现代数学基础丛书 129

概率论基础

（第二版）

严士健 王隽骧 刘秀芳 著



科学出版社
www.sciencep.com

(O—3630.0101)


ISBN 978-7-03-025155-8
9 787030 251558 >

销售分类建议：高等数学

定 价：66.00 元

现代数学基础丛书 129

概率论基础

（第二版）

严士健　王隽骧　刘秀芳　著

科学出版社
北京

内 容 简 介

本书用测度论的观点论述概率论的基本概念，如概率、随机变量与分布函数、数学期望与条件数学期望和中心极限定理等．本书特点是把测度论的基本内容与概率论的基本内容结合在一起讲述，论述严谨，条理清楚，便于自学．凡学过概率论基础课的读者都能阅读本书．每节后面附有习题，以便加深理解书中的内容．

读者对象是大学数学系高年级学生、研究生、教师及科学工作者．

图书在版编目(CIP)数据

概率论基础/严士健，王隽骧，刘秀芳著．-2版．—北京：科学出版社，2009

(现代数学基础丛书；129)

ISBN 978-7-03-025155-8

I. 概… II. ①严…②王…③刘… III. 概率论 IV. O211

中国版本图书馆 CIP 数据核字(2009) 第 134236 号

科学出版社 出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

铭浩彩色印装有限公司 印刷

科学出版社发行 各地新华书店经销

*

2009 年 8 月第 一 版 开本：B5(720 × 1000)

2009 年 8 月第一次印刷 印张：26 3/4

印数：1—2 500 字数：524 000

定价：66.00 元

(如有印装质量问题，我社负责调换〈环伟〉)

《现代数学基础丛书》序

对于数学研究与培养青年数学人才而言，书籍与期刊起着特殊重要的作用．许多成就卓越的数学家在青年时代都曾钻研或参考过一些优秀书籍，从中汲取营养，获得教益．

20 世纪 70 年代后期，我国的数学研究与数学书刊的出版由于“文化大革命”的浩劫已经被破坏与中断了 10 余年，而在这期间国际上数学研究却在迅猛地发展着．1978 年以后，我国青年学子重新获得了学习、钻研与深造的机会．当时他们的参考书籍大多还是 50 年代甚至更早期的著述．据此，科学出版社陆续推出了多套数学丛书，其中《纯粹数学与应用数学专著》丛书与《现代数学基础丛书》更为突出，前者出版约 40 卷，后者则逾 80 卷．它们质量甚高，影响颇大，对我国数学研究、交流与人才培养发挥了显著效用．

《现代数学基础丛书》的宗旨是面向大学数学专业的高年级学生、研究生以及青年学者，针对一些重要的数学领域与研究方向，作较系统的介绍．既注意该领域的基础知识，又反映其新发展，力求深入浅出，简明扼要，注重创新．

近年来，数学在各门科学、高新技术、经济、管理等方面取得了更加广泛与深入的应用，还形成了一些交叉学科．我们希望这套丛书的内容由基础数学拓展到应用数学、计算数学以及数学交叉学科的各个领域．

这套丛书得到了许多数学家长期的大力支持，编辑人员也为其付出了艰辛的劳动．它获得了广大读者的喜爱．我们诚挚地希望大家更加关心与支持它的发展，使它越办越好，为我国数学研究与教育水平的进一步提高做出贡献．

杨　乐

2003 年 8 月

再版前言

本书自 1982 年出版以来, 承蒙读者厚爱, 先后印刷五次, 三万二千余册. 这期间, 概率统计界的朋友及使用本书的读者, 不断提供了一些修改意见. 特别是安徽师范大学丁万鼎教授指出, 本书原第 1 章 6.3 节定理 1(独立类扩张定理) 的证明可以大大简化并将他的证明寄给作者; 徐州师范大学的戴朝寿教授把他在教学过程中发现的整个书中的纰漏和印刷错误寄给作者, 北京师范大学的张余辉教授也在教学过程中特别注意了对本书的修正.

这次, 应科学出版社之邀对本书进行修订. 我们对全书进行了仔细的审阅, 采纳了上述同志们的意见, 对原书进行了适当的修改和补充. 特别对于独立类扩张定理和无穷乘积概率存在定理给出了新的阐述和证明.

本次修订后仍可能有错误和不妥之处, 敬请广大读者提出宝贵意见.

作　者

2009 年 2 月

序　言

概率论是从数量上研究随机现象的规律性的学科. 它在自然科学、技术科学、社会科学和管理科学中有着广泛的应用. 因此从 20 世纪 30 年代以来, 发展甚为迅速, 新的分支学科不断涌现, 成为近代数学的一个重要组成部分.

概率论与数理统计的严格理论是以测度论为基础的. 为了适应这方面需要, 曾经 (或将要) 出版一些测度论的书籍. 这当然都是很必要的. 但是在我们的教学实践中深深感到, 有些学过概率论与数理统计基础课而想继续自学深造的读者, 在学了测度论以后, 常常不能运用自如地将它用来处理概率论的基础问题. 还有, 一些有关多维随机变量的内容很多书籍认为和一维处理方法类似, 只是述而不证, 实际上在某些重要的细节方面还是有所不同的, 读者在这些方面也难于获得参考材料. 因此, 我们认为, 为了阐明这些问题并使学生在学习测度论时能更好地明确目的性, 编写一本用测度论的观点论述概率论的书, 对概率论专业的教师和学生来说, 也许是有参考意义的.

本书的目的在于向那些已学过相当于大学概率论基础课而希望进一步学习严格数学理论的读者提供一个基础读物. 它包括了近代概率论与数理统计所必需的测度论内容、近代概率论的基本概念、工具及其性质的严格论述, 特别是条件概率、条件数学期望、条件分布及多元特征函数, 其中有些内容可能是新的. 此外在保持测度论的系统的同时, 还特别注意了应用测度论来严格处理大学概率论基础课中所没有讲清楚的问题. 对于多元随机变量的分布、特征函数及其有关性质也作了严格的处理. 本书的预备知识是大学数学系的数学分析、高等代数及复变函数论. 为便于自学, 本书对概念的阐释和推理的叙述都比较详细.

本书的初稿是 60 年代我和王隽骧同志讲授概率选修课时合写的讲义的一部分. 这次由我和刘秀芳同志进行了讨论, 由刘秀芳同志执笔, 对讲义进行全面的整理与修改, 并在全国师范院校概率统计学术会议以及北京师范大学 79 届的概率论研究生及进修教师班上讲授过.

作者们深深感谢张禾瑞教授、王梓坤教授、孙永生教授和侯振挺教授的鼓励与关心; 感谢陈木法同志以及教研室其他同志对讲义的修改、整理工作的支持和关心. 感谢北京师范大学 78、79 届的概率论专业研究生和进修教师, 他们阅读了大部分修改稿, 对习题进行了详细的检查, 提出了不少中肯的意见. 感谢徐承彝、朱作

宾二位同志详细地阅读了大部分手稿, 提出了很多宝贵的修改意见.

由于作者们的水平所限, 书中一定存在着不少缺点和错误, 恳请批评指正.

严士健

1980 年 6 月于北京师范大学

目　录

第 1 章　概率与测度

本章将在回顾概率概念的实际背景的基础上, 给出概率与测度的定义; 讨论今后常用到的一些集类 (半集代数、集代数、σ-代数等) 的基本性质; 讨论测度的性质及测度扩张问题; 最后讨论独立事件类的扩张问题.

§1.1　引　　言

概率论是研究随机现象中的数量规律的科学.

在各种自然科学 (包括数学) 中, 大部分现象的规律是以下列形式表达的: "只要条件 $\mathscr{C}$ 一经实现, 则事件 A 必然发生 (或必然不发生)." 例如, "如果平面图形是三角形 (条件 $\mathscr{C}$ 实现), 那么这个图形的内角和是180°(事件 A 一定发生)"; 又如 "一个力作用于一物体时 (条件 $\mathscr{C}$ 实现), 该物体必产生加速度 (事件 A 发生)"; 再如 "在一个标准大气压、温度 100°C 的条件下, 水一定沸腾." 以下称在条件 $\mathscr{C}$ 下必然发生的事件 A 为**条件 $\mathscr{C}$ 下的必然事件**(或简称必然事件), 称在条件 $\mathscr{C}$ 下必然不发生的事件 A 为**条件 $\mathscr{C}$ 下的不可能事件**(或简称不可能事件).

与必然事件不同, 在客观世界中还存在这样的事件: 在条件 $\mathscr{C}$ 下, 事件 A 可能发生也可能不发生. 我们以后将在条件 $\mathscr{C}$ 实现时, 可能发生也可能不发生的事件叫做**条件 $\mathscr{C}$ 下的随机事件**(或简称随机事件). 下面我们举一些例子.

例 1　从某工厂的某种产品中抽出的一件产品可能是合格品, 也可能不是. 在这个例子中, 条件 $\mathscr{C}$ 是 "从某工厂的某种产品中抽出一件产品", 事件 A 是 "抽得的产品是合格品", 显然 A 是条件 $\mathscr{C}$ 下的随机事件. 同样, 将 "从某工厂的某种产品中抽得 n 件产品" 看作条件 $\mathscr{C}$, 则 "其中恰有 k 件合格品"$(0 \leqslant k \leqslant n)$ 是在 $\mathscr{C}$ 之下的随机事件.

例 2　在无线电中继通讯中, 由于种种原因, 有时一个或几个中继站工作失效, 这时通讯就中断了. 这种通讯中断现象通常是具有偶然性的. 因此 "通讯中断" 是 "无线电中继通讯" 这一条件下的随机事件. 若有 n 个中继站, $0 \leqslant k \leqslant n$, 则 "恰有 k 个中继站工作失效" 是 "n 个中继站的无线电中继通讯" 条件下的随机事件.

例 3　对于一个公用电话系统来说, 电话局收到用户的呼叫是具有偶然性的, 因此, "在时间间隔 $(a, a+t]$ 内, 电话局收到用户的 k 次呼叫" 是 "这个公用电话系统在 $(a, a+t]$ 内工作" 的条件下的随机事件.

例 4　铀的放射性原子经过放出粒子而变成其他元素 (物理学称为蜕变), 但

是每个放射性原子什么时候发生蜕变是具有偶然性的. 因此, “在时间间隔 $(a, a+t]$ 内, 给定的一块放射性铀”(这是条件 $\mathscr{C}$)“恰有 n 个原子蜕变” 是一个随机事件.

例 5　(a) 某一射手向一目标进行一次或多次射击时, 可能击中目标, 也可能没有击中目标, 因此 “射手击中目标” 是 “该射手向目标进行一次射击” 条件下的随机事件, 也可以是 “该射手向目标进行 n 次射击” 条件下的随机事件.

(b)“弹着点距目标的偏差为 x” 及 “偏差在 x_1, x_2 之间” 都是 “进行一次射击” 条件下的随机事件. “n 个弹着点与目标的偏差恰有 k 个小于 x” 是 “进行 n 次射击” 条件下的随机事件.

(c) 如果在目标所在的平面上建立坐标系 (不妨设目标位于坐标原点), 则 “弹着点 (ξ, η) 位于区域 $\{(x,y): x_1 \leqslant x < x_2, y_1 \leqslant y < y_2\}$ 内” 是 “进行一次射击” 的条件下的随机事件.

例 6　“一个在液体中悬浮着的质点 (例如花粉) 在时刻 t 的位置 (ξ, η, ζ) 位于区域 $\{(x,y,z)$: $x_1 \leqslant x < x_2, y_1 \leqslant y < y_2, z_1 \leqslant z < z_2\}$ 内” 是随机事件.

下面我们引入概率的概念并说明概率的直观含义.

研究随机事件有两种途径: 一种自然的想法就是像我们研究必然事件那样, 去进一步寻求随机事件发生的条件, 但是只要对上面举出的一些例子分析一下, 就可以发现这种做法几乎是不可能的, 而且也是不必要的. 例如对于例 3, 要想事先找出 “在 $(a, a+t]$ 内, 电话局收到用户的 k 次呼叫” 这一随机事件发生的条件是不可能的. 另一方面, 电话局所关心的只是平均或大致有多少次呼叫, 从而估计电话局的线路及接线员是否够用, 是否有浪费现象. 因此, 寻求事件确切发生的条件的要求也是不必要的. 既然电话局所关心的问题是 “大致” 有多少次呼叫, 因此就产生了求这个随机事件出现的 “可能性” 的想法.

这种寻求随机事件 “发生的可能性” 的想法, 在人们实践中是很自然的. 例如, 人们对于工厂生产好坏的判别标准之一是它的产品的合格率, 即产品中合格品个数与产品总数的比值; 一种药物对某种疾病的疗效用治愈率, 即使用过该药的患者中治愈人数与总人数之比来判断; 对于射手的射击技术的判别标准之一是它的命中率, 即射击命中次数与射击总次数之比. 合格率、治愈率及命中率都是人们所关心的随机事件发生的可能性的一种描述. 从这些例子可以概括出频数与频率的概念如下: 若在 $\mathscr{C}$ 的 n 次实现之下, 事件 A 发生的次数是 μ_A, 称它为 A 在 $\mathscr{C}$ n 次实现之下的**频数**; 而称 $\frac{\mu_A}{n}$ 为事件 A 在 $\mathscr{C}$ n 次实现之下的**频率**(简称频率), 记作 ν_A, 它表示 A 在 $\mathscr{C}$ 之下发生的可能性. 当进一步考察 A 的频率 ν_A 时, 我们发现 ν_A 对于 $\mathscr{C}$ 的不同的 n 次实现一般来说是不同的, 但是大量的实践表明, 对于固定的随机事件 A 来说, 当 n(条件 $\mathscr{C}$ 实现的次数) 较大时, ν_A 有经常接近于一个常数的趋势, 并且当 n 越大时, 接近的程度也就更为显著, 接近的次数也越经常. 因此有

理由认为这个与 A 有关的常数 (记作 $P(A)$) 刻划了 A 的一种重要特性 —— 事件 A 发生的可能性. 我们称 $P(A)$ 为**随机事件 A 在条件 $\mathscr{C}$ 下的概率.**

实际上, 上面提出的概率概念只是一个说明, 并不能作为数学定义. 为了将概率论建立在严格的基础上, 首先要对随机事件给出数学定义, 然后给出概率的定义. 下面几节将逐步进行这一工作.

§1.2 事件与集合

本节目的是在对随机事件进一步分析的基础上, 将它定义成某一与条件有关的集 (基本事件空间)的子集, 然后引进事件运算的概念.

1.2.1 基本事件空间与事件

从上一节例子可以看出: 在给定条件 $\mathscr{C}$ 下, 随机事件通常有很多个. 我们现在来研究这些随机事件能否由一组 "基本" 事件 "组成" 的问题. 先看一些例子.

例 1 设 $\mathscr{C}$ 表示 "从某工厂的某种产品中抽出 n 件" 这一条件. 如果我们感兴趣的是研究在 $\mathscr{C}$ 下抽得产品中的合格品件数的问题, 那么很自然地要考虑随机事件 $A_k, k = 0, 1, \cdots, n$, 其中 A_k 表示 "抽得的产品中合格品的件数为 k"; 一般, 可以考虑合格品的件数在某一范围之内的事件, 例如 $\widetilde{B}_k$——"抽得产品中合格品的件数不超过 k" 这类事件; 更一般的情况是考虑事件 $\widetilde{F}_\Lambda$, 它表示事件 "抽得产品中合格品的件数 $k \in \Lambda$", 其中 Λ 是整数集 $Z_n \triangleq \{0, 1, 2, \cdots, n\}$①的任一子集. 也就是说, 若 $\Lambda = \{k_1, \cdots, k_m\} \subset Z_n$, 则 $\widetilde{F}_\Lambda$ 表示事件 "抽得产品中合格品的件数为 k_1, 或 $k_2, \cdots$, 或 k_m".

现在我们来分析一下事件 $A_k, \widetilde{B}_k$ 及 $\widetilde{F}_\Lambda$ 之间的关系. 容易看到, $\widetilde{B}_k$ 也可以表达成 "抽得产品中合格品的件数是 0, 或是 $1, \cdots$, 或者 k". 于是在条件 $\mathscr{C}$ 的每一次实现之下 (即每 "抽 n 件产品" 一次), 事件 $\widetilde{B}_k$ 发生当且仅当 $A_0, A_1, \cdots, A_k$ 有一个发生. 这说明事件 $\widetilde{B}_k$ 与事件集 $\{A_0, A_1, \cdots, A_k\}$ 相互唯一决定 (实际上 $\widetilde{B}_k$ 可以表达成 "A_0, 或 $A_1, \cdots$, 或A_k"). 一般地, 由上一段对 $\widetilde{F}_\Lambda$ 的解释可知: 在条件 $\mathscr{C}$ 的每一次实现之下, 事件 $\widetilde{F}_\Lambda$ 发生当且仅当 $A_{k_1}, \cdots, A_{k_m}$ 有一发生, 从而事件 $\widetilde{F}_\Lambda$ 与事件集 $\{A_{k_1}, \cdots, A_{k_m}\}$ 相互唯一决定. 这样, 我们就可以把事件 $\widetilde{F}_\Lambda$ 与事件集 $\{A_k : k \in \Lambda\}$ 等同起来. 又 $\widetilde{F}_{Z_n}$ 显然是 $\mathscr{C}$ 之下的必然事件, 按照上述约定, 它与事件集 $\{A_k : k \in Z_n\}$ 等同. 另一方面, 在 $\mathscr{C}$ 的每一次实现之下, 事件 $A_0, A_1, \cdots, A_n$ 有一且仅有一发生. 因此从考虑 $\mathscr{C}$ 之下合格品件数问题的角度来看, 事件 $A_0, A_1, \cdots, A_n$ 可以唯一地表示一切事件 $\widetilde{F}_\Lambda, \Lambda \subset Z_n$; (同时它本身也 "不必再分细") 从而可以认为: 事件 $A_0, A_1, \cdots, A_n$ 是全体基本事件; 事件集

① 符号 $\triangleq$ 读作 "定义作".

$\{A_0, A_1, \cdots, A_n\}$ 是基本事件空间, 记作 Ω; 它本身表示在条件 $\mathscr{C}$ 下的必然事件; $\widetilde{F}_\Lambda$ 定义为 Ω 的子集 $F_\Lambda \hat{=} \{A_k : k \in \Lambda\}$.

例 2 (参看 §1.1 例 3)　用 $\mathscr{C}$ 表示 "某公用电话系统在时间间隔 $(a, a+t]$ 内工作" 这一条件. 假定我们希望研究的是总机在 $(a, a+t]$ 内收到用户的呼叫次数问题. 那么和例 1 类似, 可以只考虑事件 $\widetilde{F}_\Lambda$: 现在 $\widetilde{F}_\Lambda$ 表示 "总机在 $(a, a+t]$ 内收到用户的呼叫次数 $k \in \Lambda$", 其中 Λ 是非负整数集 $Z_+ = \{0, 1, 2, \cdots\}$ 的任一子集.

令 $A_k = \widetilde{F}_{\{k\}}, k \in Z_+$, 和例 1 的分析完全类似 (建议读者仔细分析一下) 可以得到下述结论: 在条件 $\mathscr{C}$ 下, 考虑总机收到用户呼叫次数的问题时, 可以认为事件集 $\Omega = \{A_k, k \in Z_+\}$ 是基本事件空间, 它本身表示在条件 $\mathscr{C}$ 下的必然事件, $\widetilde{F}_\Lambda$ 定义为 Ω 的子集 $F_\Lambda \hat{=} \{A_k : k \in \Lambda\}$.

例 3 (参看 §1.1 例 5)　用 $\mathscr{C}$ 表示 "某射手向目标进行一次射击" 这一条件. 如果我们希望研究弹着点与目标的偏差问题, 那么一般可以只考虑这样一些随机事件: $\widetilde{B}_I$——"弹着点与目标的偏差 $x \in I$", 其中 I 表示非负实半轴 $R_+ \hat{=} [0, \infty)$ 的任一子区间.

令 $\omega_x, x \in R_+$ 表示随机事件 "弹着点与目标的偏差为 x", 和例 1 的分析类似可以得出结论: $\omega_x, x \in R_+$, 是基本事件, 集 $\Omega = \{\omega_x, x \in R_+\}$ 是基本事件空间, $\widetilde{B}_I, I \subset R_+$, 定义为 Ω 的子集 $B_I \hat{=} \{\omega_x : x \in I\}$.

例 4 (参看 §1.1 例 6)　用 $\mathscr{C}$ 表示 "液体中有一悬浮质点" 这一条件, 我们希望研究这一质点在时刻 t 的位置问题. 可以只考虑随机事件 $\widetilde{B}_D =$ "$(\xi, \eta, \zeta) \in D$", 其中 D 表示三维直角坐标空间 $R^{(3)}$ 的某一区域 (例如不妨先只考虑 D 是可求体积的区域), (ξ, η, ζ) 表示该悬浮质点在时刻 t 的位置.

以 $\omega_{(x,y,z)}, (x, y, z) \in R^{(3)}$, 表示随机事件 "$(\xi, \eta, \zeta) = (x, y, z)$" 即 "悬浮质点在时刻 t 的位置在 (x, y, z) 处". 那么和前面的分析类似, 可以认为 $\Omega = \{\omega_{(x,y,z)} : (x, y, z) \in R^{(3)}\}$ 为基本事件空间, Ω 表示在条件 $\mathscr{C}$ 下的必然事件. 而 $\widetilde{B}_D$ 定义为 $B_D \hat{=} \{\omega_{(x,y,z)} : (x, y, z) \in D\}$.

这些例子说明, 当我们希望研究某随机事件类 $\widetilde{\mathfrak{C}}$ 时, 可以找到一族在条件 $\mathscr{C}$ 下 "不能或不必再分"(对 $\widetilde{\mathfrak{C}}$ 来说) 的随机事件, 它们组成的集合记作 Ω, 使得在 $\mathscr{C}$ 的每次实现之下有一且仅有一 $\omega \in \Omega$ 发生; 而且随机事件类 $\widetilde{\mathfrak{C}}$ 与 Ω 的某一子集类 $\mathfrak{C}$ 按下述法则一一对应: $\widetilde{A} \in \widetilde{\mathfrak{C}}$ 与 $A \in \mathfrak{C}$ 对应是指在 $\mathscr{C}$ 的每次实现之下, $\widetilde{A}$ 发生当且仅当有一 $\omega \in A$ 发生. 根据这个法则, 我们今后将事件 $\widetilde{A}$ 与 A 等同, 而且认为研究条件 $\mathscr{C}$ 下的随机事件类 $\widetilde{\mathfrak{C}}$, 就是研究相应的集 Ω 的子集类 $\mathfrak{C}$, 而 Ω 称为 $\mathscr{C}$ 下的**基本事件空间**, $\mathfrak{C}$ 中的元 (Ω 的子集) 称为 (随机)**事件**, Ω 本身是**必然事件**, 空集 $\varnothing$ 是**不可能事件**.

这里我们要强调指出 "不能或不必再分" 是相对于所考虑的随机事件类而言, 并不是绝对的. 事实上, 在给定条件 $\mathscr{C}$ 下, 基本事件和基本事件空间的选取并不是

唯一的, 而且对具体问题来说, 基本事件选取恰当, 常常有助于更快地解决问题.

例 5 在例 1 中, 我们也可以选取另外的事件为基本事件. 我们用 ω 表示 "抽出一件产品是合格品" 这一事件, 用 ω^c 表示 "抽出一件产品不是合格品" 这一事件; 然后用 n 元组 $(\omega_1, \omega_2, \cdots, \omega_n)$, 其中每一 ω_k 可取 ω 或 ω^c, 表示依次抽出 n 件产品所得到的随机事件. 这样, 每抽出 n 件产品, 下列 2^n 个随机事件

$$(\omega_1, \omega_2, \cdots, \omega_n), \quad \omega_m = \omega \text{ 或 } \omega^c, \quad m = 1, 2, \cdots, n. \tag{1}$$

有一且仅有一发生. 于是由 (1) 中所列出的 2^n 个随机事件作成的集就是基本事件空间. 此时例 1 中的事件 A_k 在目前情况下便不再是基本事件, 而是 (1) 中恰有 k 个 $\omega_m = \omega$ 的那些基本事件组成的集. 这样的基本事件共有 $\binom{n}{k}$ 个, 其中 $\binom{n}{k}$ 表示 n 个不同元素中取 k 个的组合数.

例 6 在例 3 中, 有时也可以选另外的事件. 我们采取 § 1.1 例 5(c) 的办法. 弹着点的坐标以 (ξ, η) 表示, 以 $\lambda_{(x,y)}$ 表示随机事件 "$(\xi, \eta) = (x, y)$", $(x, y) \in R^{(2)}$(二维直角坐标空间), 于是在 "进行一次射击" 的条件下, $\lambda_{(x,y)}, (x, y) \in R^{(2)}$, 有且仅有一发生, 从而随机事件集 $\Omega_1 = \{\lambda_{(x,y)}, (x, y) \in R^{(2)}\}$ 也可以认为是 "进行一次射击" 的条件下的基本事件空间, 并且例 3 中的随机事件 $\omega_z, z \in R_+$, 在 Ω_1 中应该是 $\{\lambda_{(x,y)} : x^2 + y^2 = z^2\}$, 而 $\widetilde{B}_I$, 应该是 Ω_1 中的子集 $\{\lambda_{(x,y)} : \sqrt{x^2 + y^2} \in I\}$.

1.2.2 事件的运算与集的运算

为使概率能作为数学对象运算, 必须要讨论随机事件之间的关系. 在给出数学定义之前, 先作一些直观考察.

设 $\widetilde{A}, \widetilde{B}, \widetilde{C}$ 是在条件 $\mathscr{C}$ 之下的三个随机事件, 若在 $\mathscr{C}$ 的每次实现下, $\widetilde{C}$ 发生当且仅当 $\widetilde{A}, \widetilde{B}$ 至少有一发生, 则称 $\widetilde{C}$ 为 $\widetilde{A}, \widetilde{B}$**的和**, 记作 $\widetilde{C} = \widetilde{A} \cup \widetilde{B}$; 若在 $\mathscr{C}$ 的每次实现下, $\widetilde{C}$ 发生当且仅当 $\widetilde{A}, \widetilde{B}$ 都发生, 则称 $\widetilde{C}$ 为 $\widetilde{A}, \widetilde{B}$**的积**, 记作 $\widetilde{C} = \widetilde{A} \cap \widetilde{B}$; 若在 $\mathscr{C}$ 的每次实现之下, $\widetilde{A}$ 发生导致 $\widetilde{B}$ 发生, 则称 $\widetilde{B}$**包含**$\widetilde{A}$, 记作 $\widetilde{A} \subset \widetilde{B}$ 或 $\widetilde{B} \supset \widetilde{A}$. 例如, 若 $\widetilde{A}, \widetilde{B}$ 分别是例 1 中的 $\widetilde{F}_{\Lambda_1}, \widetilde{F}_{\Lambda_2}$, 则 $\widetilde{F}_{\Lambda_1} \cup \widetilde{F}_{\Lambda_2} = \widetilde{F}_{\Lambda_1 \cup \Lambda_2}, \widetilde{F}_{\Lambda_1} \cap \widetilde{F}_{\Lambda_2} = \widetilde{F}_{\Lambda_1 \cap \Lambda_2}$ 且若 $\Lambda_1 \subset \Lambda_2$, 则 $\widetilde{F}_{\Lambda_1} \subset \widetilde{F}_{\Lambda_2}$, 其中 $\Lambda_1 \cup \Lambda_2, \Lambda_1 \cap \Lambda_2, \Lambda_1 \subset \Lambda_2$ 分别表示集 Λ_1, Λ_2 的并、交及包含关系. 显然容易从例 2—4 举出各种具体例子.

因此按照 1.2.1 的法则易证, 在 $\mathscr{C}$ 的每一次实现之下, $\widetilde{A} \cup \widetilde{B}$(相应地: $\widetilde{A} \cap \widetilde{B}$) 发生当且仅当有一 $\omega \in A$ 或 $\omega \in B$(相应地: $\omega \in A$ 且 $\omega \in B$) 发生, 即有一 $\omega \in A \cup B$(相应地: $\omega \in A \cap B$) 发生 (其中 $A \cup B, A \cap B$ 分别表示集 A, B 的并与交). 因此, 事件的和与积分别与 Ω 的相应子集的并与交对应. 其次, 若 $\widetilde{A} \subset \widetilde{B}$, 则在 $\mathscr{C}$ 的每次实现下, $\widetilde{A}$ 发生当且仅当 $\widetilde{A}, \widetilde{B}$ 都发生, 即当且仅当 $\widetilde{A} \cap \widetilde{B}$ 发生, 所以有 $\widetilde{A} = \widetilde{A} \cap \widetilde{B}$. 按照 1.2.1 的对应规则及上面证明的 $\widetilde{A} \cap \widetilde{B}$ 与 $A \cap B$ 对应的结论知 $A = A \cap B$, 即

$A \subset B$, 这就证明了: 若 $\widetilde{A} \subset \widetilde{B}$, 则 $A \subset B$. 上面各个证明步骤都是可以反推的, 所以有, 若 $A \subset B$, 则 $\widetilde{A} \subset \widetilde{B}$. 因此, 事件的包含关系 $\widetilde{A} \subset \widetilde{B}$ 等价于 Ω 中相应子集 A, B 的包含关系 $A \subset B$.

定义 1　设 Ω 是一集合, 称为一个基本事件空间, $\mathfrak{E}$ 是 Ω 的某些子集作成的类, 称为关于 Ω 的**事件类**, 对于 $A, B \in \mathfrak{E}$ 若 $A \cup B \in \mathfrak{E}$(相应地: $A \cap B \in \mathfrak{E}$) 则称 $A \cup B$(相应地: $A \cap B$) 为事件 A, B 的**并或和**(相应地: **交或积**). 若 $\Omega \in \mathfrak{E}$, 则称 Ω 为**必然事件**; 若 $\varnothing \in \mathfrak{E}$, 则称 $\varnothing$ 为**不可能事件**. 对 $A, B \in \mathfrak{E}$, 若 $A \cap B = \varnothing$, 则称 A, B**互不相容**; 若 $A \cap B = \varnothing$ 且 $A \cup B = \Omega$ 则称 A, B 互为**对立事件**, A 的对立事件 (余集) 记作 A^c: 事件 $A - B = A \cap B^c$ 称为 A, B 的**差**, $A \triangle B = (A - B) \cup (B - A)$ 称为 A, B 的**对称差**; 若 $A \subset B$, 则称事件A **被** B **包含或**B **包含** A.

由以上定义可以看出随机事件类就是一个集 Ω 的一个子集类. 这样定义使我们可以充分应用集合论的结论和公式. 但对于具体的随机现象, 还需要具体决定 Ω 及 $\mathfrak{E}$.

为了应用, 我们还将事件的运算推广到事件的有限类或可数类上去, 这就是

定义 2　设 $\mathfrak{E}$ 是关于基本事件空间 Ω 的一个事件类, 若 $\{A_n : n \in I\} \subset \mathfrak{E}$, 其中 I 是非空有限集或可数集, 则 $\bigcup\limits_{n \in I} A_n$, $\bigcap\limits_{n \in I} A_n$(若它们仍属于 $\mathfrak{E}$ 的话) 分别称为 $\{A_n : n \in I\}$ 的**并**, **交**; 若 $I = \varnothing$, 则约定 $\bigcup\limits_{n \in I} A_n = \varnothing$, $\bigcap\limits_{n \in I} A_n = \Omega$; 若$A_n, n \in I$ 两两不交, 则称 $\{A_n, n \in I\}$**两两互不相容**.此时称它们的并为**和**, 记作 $\sum\limits_{n \in I} A_n$.

下面列出一些事件的运算性质 (即集合运算性质):

(i) $(A \cup B) \cup C = A \cup (B \cup C)$;

$(A \cap B) \cap C = A \cap (B \cap C)$;

$(A \triangle B) \triangle C = A \triangle (B \triangle C)$;

(ii) $A \cup B = B \cup A$;　$A \cap B = B \cap A$;　$A \triangle B = B \triangle A$;

(iii) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$;

$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$;

$(A \triangle B) \cap C = (A \cap C) \triangle (B \cap C)$;

(iv) $A \cap A^c = \varnothing$;　$A \cup A^c = \Omega$;

(v) $(A^c)^c = A$;　$\Omega^c = \varnothing$;　$\varnothing^c = \Omega$;

(vi) $\left(\bigcup\limits_{n \in I} A_n\right)^c = \bigcap\limits_{n \in I} A_n^c$;　$\left(\bigcap\limits_{n \in I} A_n\right)^c = \bigcup\limits_{n \in I} A_n^c$.

习题及补充

1. 试证:

i) $A\triangle B=(A\cup B)-(A\cap B)=(A\cap B^c)\cup(A^c\cap B)$;

$A\triangle B=B\triangle A$;

ii) $A\triangle\varnothing=A; A\triangle\Omega=A^c$; $A\triangle B=A^c\triangle B^c$;

iii) $(A\triangle B)\triangle C=A\triangle(B\triangle C)$;

$(A\triangle B)\cap C=(A\cap C)\triangle(B\cap C)$;

iv) $A\cup B=(A\triangle B)+(A\cap B)$;

v) $(A_1\cup A_2)\triangle(B_1\cup B_2)\subset(A_1\triangle B_1)\cup(A_2\triangle B_2)$;

vi) $(A_1-A_2)\triangle(B_1-B_2)\subset(A_1\triangle B_1)\cup(A_2\triangle B_2)$.

2. 若 $N\subset B$, 试证:

i) $A\cup N=(A-B)\triangle[B\cap(A\cup N)]$;

ii) $A\triangle N=(A-B)\cup[B\cap(A\triangle N)]$.

3. 设 $A_1,A_2,\cdots$ 是任一集序列, 若令

$$A_n'=A_n-\left(\bigcup_{k=1}^{n-1}A_k\right),\qquad n=1,2,\cdots,$$

则 $A_1',A_2',\cdots$ 两两不相交且

$$\bigcup_{n=1}^{\infty}A_n=\sum_{n=1}^{\infty}A_n'.$$

4. 设 $\{A_n\}$ 是一个集序列, 若 A^* 是由一切具有下列性质的元素 ω 作成的集: ω 属于无穷多个 A_n, 则称 A^* 为**集序列$\{A_n\}$的上极限**, 并记作

$$A^*=\varlimsup_{n\to\infty}A_n\qquad\left(\text{或}\ \limsup_{n\to\infty}A_n\right)$$

若 A_* 是由一切具有下列性质的元素 ω 作成的集: 只有有限个 A_n 不含有 ω, 则称 A_* 为**集序列$\{A_n\}$的下极限**, 并记作

$$A_*=\varliminf_{n\to\infty}A_n(\text{或}\liminf_{n\to\infty}A_n).$$

试证:

i) $A_*=\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty}A_m,\qquad A^*=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}A_m,\qquad A_*\subset A^*$;

ii) 若 $A_1\subset A_2\subset\cdots$, 则 $A_*=A^*=\bigcup_{n=1}^{\infty}A_n$;

若 $A_1 \supset A_2 \supset \cdots$,　则 $A_* = A^* = \bigcap_{n=1}^{\infty} A_n$;

iii) 若 A, B 是任意两个给定的集合, 且

$$A_n = \begin{cases} A, & \text{当 } n \text{ 为偶数}, \\ B, & \text{当 } n \text{ 为奇数}, \end{cases}$$

则 $\varliminf_{n\to\infty} A_n = A \cap B$,　$\varlimsup_{n\to\infty} A_n = A \cup B$;

iv) 若 A 是任一集合, 则有

$$A - A_* = \varlimsup_{n\to\infty} (A - A_n);$$
$$A - A^* = \varliminf_{n\to\infty} (A - A_n).$$

设 $A_n = (-\infty, x_n), n = 1, 2 \cdots, x_n$ 是实数, 试求集合 A_*, A^*.

§1.3　集类与单调类定理

在对随机现象的研究中, 我们要通过事件的相互联系去认识随机现象的规律性, 因此往往是要研究一下事件类, 这个事件类要在某些运算之下封闭. 我们已经建立了事件的集合表示, 因此我们需要研究对某些运算封闭的集类.

1.3.1　半集代数与集代数

半集代数 (亦称半布尔代数、具单位的半环), 是我们经常遇到的一种集类, 它是从直线上的区间类、平面上的矩形类抽象出来的一种集类.

定义 1　Ω 的子集类 $\mathscr{S}$ 称为 Ω 中的**半集代数**, 如果它满足

I) $\Omega \in \mathscr{S}, \varnothing \in \mathscr{S}$;

II) 若 $A, B \in \mathscr{S}$, 则 $A \cap B \in \mathscr{S}$;

III) 若 $A, A_1 \in \mathscr{S}, A_1 \subset A$, 则存在 n, 存在 $A_2, \cdots, A_n \in \mathscr{S}, A_1, \cdots, A_n$ 两两不交且 $A = \sum_{k=1}^{n} A_k$.

性质 1　在 I, II 成立的条件下, III 等价于

III′) 若 $A \in \mathscr{S}$, 则存在 n, 存在 $A_1, \cdots, A_n \in \mathscr{S}$, 两两不交且 $A^c = \sum_{k=1}^{n} A_k$.

证　若 III 成立, 设 $A \in \mathscr{S}$, 则 $A \subset \Omega \in \mathscr{S}$, 因而存在 $A_1, \cdots, A_n \in \mathscr{S}, A, A_1, \cdots, A_n$ 两两不交且 $\Omega = A + \sum_{k=1}^{n} A_k$, 于是 $A^c = \sum_{k=1}^{n} A_k$, 即 III′ 成立. 反之, 若 III′ 成立, 设 $A_1, A \in \mathscr{S}$ 且 $A_1 \subset A$, 则存在 $A_2', \cdots, A_n' \in \mathscr{S}$, 两两不交且

$A_1^c = \sum\limits_{k=2}^{n} A'_k$, 则 $A = A \cap A_1 + A \cap A_1^c = A_1 + \sum\limits_{k=2}^{n}(A'_k \cap A)$, 令 $A_k = A'_k \cap A$, 则由 II 知 $A_k \in \mathscr{S}, k = 2, \cdots, n$ 且 $A_1, \cdots, A_n$ 两两不交, $A = \sum\limits_{k=1}^{n} A_k$.

例 1 在 §1.2 例 2 中, 若令 $\Omega = \{A_k : k \in Z_+\}, F_{ij} = \{A_k : i \leqslant k < j\}$, 则 $\mathscr{S} = \{F_{ij} : i \leqslant j, i, j \in Z_+ 或 j = \infty\}$ 是 Ω 中的半集代数.

例 2 在 §1.2 例 3 中, 令 $\mathscr{S} = \{B_I : I = [a, b), 0 \leqslant a \leqslant b \leqslant \infty\}$ 是 $\Omega \widehat{=} \{\omega_x : x \in [0, \infty)\}$ 中的半集代数.

例 3 设 $\Omega \widehat{=} R$(即$(-\infty, \infty)$), 则
$\mathscr{S}^{(1)} = \{A : A = R 或 [a, b) 或 (-\infty, b) 或 [a, \infty), a, b \in R\}$ 是 R 中的半集代数, 其中 $[a, b) = \{x \in R : a \leqslant x < b\}$.

例 4 设 $\Omega = R^{(2)}$, 令

$$\mathscr{S}^{(2)} = \{[a, b) : a = (a_1, a_2), a_1, a_2 \in R 或 = -\infty, \quad b = (b_1, b_2), b_1, b_2 \in R 或 = +\infty\},$$

其中 $[a, b) = \{(x, y) : a_1 \leqslant x < b_1, a_2 \leqslant y < b_2, 当 a_1 或 a_2 = -\infty 时, “\leqslant” 理解为 “<”\}$. 则 $\mathscr{S}^{(2)}$ 是 $R^{(2)}$ 中的半集代数.

证 $\mathscr{S}^{(2)}$ 显然满足 I, II, 故只需验证 III′ 成立. 设 $A = [a, b), a = (a_1, a_2), b = (b_1, b_2)$, 若令

$$\begin{aligned}
A_1 &= \{(x, y) : \quad -\infty < x < b_1, \quad -\infty < y < a_2\}; \\
A_2 &= \{(x, y) : \quad -\infty < x < a_1, \quad a_2 \leqslant y < +\infty\}; \\
A_3 &= \{(x, y) : \quad a_1 \leqslant x < +\infty, \quad b_2 \leqslant y < +\infty\}; \\
A_4 &= \{(x, y) : \quad b_1 \leqslant x < +\infty, \quad -\infty < y < b_2\},
\end{aligned}$$

注意: 若 a_1 或 $a_2 = -\infty$, 上述各集合的表示式中 $a_1 \leqslant$ 或 $a_2 \leqslant$ 相应地改为 “$a_1 <$” 或 “$a_2 <$”, 则 $A_1, \cdots, A_4$ 皆属于 $\mathscr{S}^{(2)}$ 且两两不交, 并有

$$A^c = R^{(2)} - A = \sum_{i=1}^{4} A_i,$$

故 III′ 成立.

我们研究随机事件, 经常需要考虑事件的和、差、积以及对立事件, 于是就需要研究对并、差、交及求余运算封闭的集类. 这就是集代数 (亦称布尔代数或具有单位的环) 的概念.

定义 2 Ω 的子集类 $\mathscr{F}$ 称为 Ω 中的**集代数**(亦称布尔代数), 如果它满足

I) $\Omega \in \mathscr{F}$;

II) 若 $A, B \in \mathscr{F}$, 则 $A \cup B, A \cap B, A - B \in \mathscr{F}$;

III) 若 $A \in \mathscr{F}$, 则 $A^c \in \mathscr{F}$.

§ 1.2 例 1、例 2 及例 4 中的集类都是集代数.

性质 2　Ω 的子集类 $\mathscr{F}$ 是集代数的充分必要条件是它满足 I, III 及

II′) 若 $A \in \mathscr{F}, B \in \mathscr{F}$, 则 $A \cup B \in \mathscr{F}$.

证　必要性是显然的. 要证充分性只需证明: 若 $A, B \in \mathscr{F}$,

则 $A - B \in \mathscr{F}, A \cap B \in \mathscr{F}$, 今证明此事.

由集合的运算可知, 对 Ω 的任何子集 A, B 有

(1) $(A \cap B)^c = A^c \cup B^c$;

(2) $(A^c)^c = A$;

(3) $A - B = A \cap B^c$

成立, 故若 $A, B \in \mathscr{F}$, 则由 III 知 $A^c \in \mathscr{F}, B^c \in \mathscr{F}$, 因而由 II′ 知 $A^c \cup B^c \in \mathscr{F}$, 故由 (1) 知 $(A \cap B)^c \in \mathscr{F}$, 再由 III 及 (2) 知

(4) $A \cap B \in \mathscr{F}$.

其次若 $A, B \in \mathscr{F}$, 则由 III 知 $A, B^c \in \mathscr{F}$, 于是由 (3), (4) 知 $A - B = A \cap B^c \in \mathscr{F}$, 因此性质 2 获证. □

性质 3　Ω 的子集类 $\mathscr{F}$ 是集代数的充分必要条件是 I, III 及

II″) 若 $A, B \in \mathscr{F}$, 则 $A \cap B \in \mathscr{F}$;

或 I 及

III′) 若 $A, B \in \mathscr{F}$, 则 $A - B \in \mathscr{F}$.

应用性质 2 及集合的运算性质不难证明此性质, 留给读者来完成.

由归纳法不难证明集代数对有限并、有限交运算封闭.

半集代数与集代数有着密切的关系, 首先, 任一个集代数显然是半集代数; 其次, 如果给了一个半集代数 $\mathscr{S}$, 不难构造一个包含 $\mathscr{S}$ 的集代数, 这就是下面的

定理 1　若 $\mathscr{S}$ 是 Ω 中的半集代数, 则集类

$$\mathscr{F} \triangleq \left\{ A : A = \sum_{k=1}^{n} A_k, n \geqslant 1, A_k \in \mathscr{S}, \quad k = 1, 2, \cdots, n, \text{两两不交.} \right\}$$

是集代数, 且 $\mathscr{F}$ 是包含 $\mathscr{S}$ 的最小集代数, 即若有集代数 $\mathscr{F}' \supset \mathscr{S}$, 则 $\mathscr{F}' \supset \mathscr{F}$. 包含 $\mathscr{S}$ 的最小集代数记作 $\mathscr{F}(\mathscr{S})$.

证　由于集代数对有限并运算封闭, 故若 $\mathscr{F}'$ 是包含 $\mathscr{S}$ 的集代数, 则必有 $\mathscr{F}' \supset \mathscr{F}$. 因此如果 $\mathscr{F}$ 是集代数, 则 $\mathscr{F}$ 必是包含 $\mathscr{S}$ 的最小集代数, 今往证 $\mathscr{F}$ 是集代数.

由于 $\Omega \in \mathscr{S} \subset \mathscr{F}$, 故由性质 3 知要证 $\mathscr{F}$ 是集代数, 只需证: 若 $A, B \in \mathscr{F}$, 则

$A\cap B, A^c\in\mathscr{F}$. 由 $A,B\in\mathscr{F}$, 有

$$A=\sum_{i=1}^{n}A_i,\quad B=\sum_{j=1}^{m}B_j,$$

$\{A_1,\cdots,A_n\}\subset\mathscr{S},\{B_1,\cdots,B_m\}\subset\mathscr{S}$ 都是两两不交的子集系,

$$A\cap B=\sum_{i=1}^{n}\sum_{j=1}^{m}A_i\cap B_j,$$

由半集代数定义知 $A_i\cap B_j\in\mathscr{S}, i=1,\cdots,n, j=1,\cdots,m$ 且两两不交, 从而 $A\cap B\in\mathscr{F}$, 即 $\mathscr{F}$ 对有限交封闭. 而

$$A^c=\bigcap_{i=1}A_i^c,$$

由性质 1 知每一 A_i^c 是 $\mathscr{S}$ 中有限个两两个交集的并, 故 $A_i^c\in\mathscr{F}$, 再由 $\mathscr{F}$ 对有限交运算封闭即知 $A^c\in\mathscr{F}$. □

例 5 令 $\mathscr{S}^{(1)}$ 如例 3 所述, 则

$$\mathscr{F}\triangleq\Big\{A:A\subset R, A=\sum_{i=1}^{n}A_i, A_i\in\mathscr{S}^{(1)},\ i=1,\cdots n \text{ 两两不交}, n\geqslant 1\Big\}$$

是一 R 中的集代数

例 6 若 Ω 是任一非空集合, $A\subset\Omega$, 令

$$\mathscr{F}=\{\varnothing,\Omega,A,A^c\}.$$

则 $\mathscr{F}$ 是一 Ω 中的集代数.

1.3.2 σ-代数

在研究随机事件时, 不仅涉及到事件的有限并、有限交, 而且往往涉及到事件的可数并与可数交, 于是需要研究对可数运算封闭的事件类.

定义 3 称 Ω 的子集类 $\mathscr{A}$ 为 Ω 中的 **σ-代数,** 如果它满足条件

I) $\Omega\in\mathscr{A}$;

II) 若 $A\in\mathscr{A}$, 则 $A^c\in\mathscr{A}$;

III) 若 $A_n\in\mathscr{A}, n=1,2,\cdots$, 则 $\bigcup_{n=1}^{\infty}A_n\in\mathscr{A}$.

性质 4 任一 σ-代数 $\mathscr{A}$ 必为一集代数.

证　由性质 2 知, 只需证明 $\mathscr{A}$ 对有限并运算封闭. 若 $A, B \in \mathscr{A}$, 则令 $A_1 = A, A_2 = B, A_n = \varnothing, n \geqslant 3$, 而 $\varnothing = \Omega^c \in \mathscr{A}$, 于是 $A_n \in \mathscr{A}, n = 1, 2, \cdots$. 由 σ-代数定义中的 III 知 $A \cup B = \bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{A}$. 故得证. □

性质 5　Ω 的子集类 $\mathscr{A}$ 是一个 σ-代数的充分与必要条件是 I, II 及

III′) 若 $A_n \in \mathscr{A}, n = 1, 2, \cdots$, 则 $\bigcap\limits_{n=1}^{\infty} A_n \in \mathscr{A}$.

证　先证必要性. 若 $\mathscr{A}$ 是 σ-代数且 $A_n \in \mathscr{A}, n = 1, 2, \cdots$, 则由 II 知 $A_n^c \in \mathscr{A}, n = 1, 2, \cdots$, 由 III 知 $\bigcup\limits_{n=1}^{\infty} A_n^c \in \mathscr{A}$, 而 $\bigcup\limits_{n=1}^{\infty} A_n^c = \left(\bigcap\limits_{n=1}^{\infty} A_n\right)^c$, 再由 II 知 $\bigcap\limits_{n=1}^{\infty} A_n = \left(\bigcup\limits_{n=1}^{\infty} A_n^c\right)^c \in \mathscr{A}$. 故 III′ 成立.

再证充分性: 若 $\mathscr{A}$ 是 Ω 的子集类, 且满足 I, II 及 III′, 今证 III 成立. 若 $A_n \in \mathscr{A}, n = 1, 2, \cdots$, 则由公式 $\left(\bigcup\limits_{n=1}^{\infty} A_n\right)^c = \bigcap\limits_{n=1}^{\infty} A_n^c$ 及上述方法易证 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{A}$. 即 III 成立. □

性质 6　Ω 的一切子集构成的集类是 Ω 中的一个 σ-代数.

易知, Ω 的一切子集构成的集类对任何集运算都是封闭的, 因此它也是半集代数、集代数、σ-代数 $\cdots$.

性质 7　任意多个 Ω 中的 σ-代数的交仍然是 Ω 中的 σ-代数. 即若 $\{\mathscr{A}_\Upsilon\}, \Upsilon \in \Gamma$, ($\Gamma$是任一非空指标集) 是 Ω 中的 σ-代数作成的类, 则 $\bigcap\limits_{\Upsilon \in \Gamma} \mathscr{A}_\Upsilon$ 是 Ω 中的一个 σ-代数.

证　令 $\mathscr{A} = \bigcap\limits_{\Upsilon \in \Gamma} \mathscr{A}_\Upsilon$. 由于 $\Omega \in \mathscr{A}_\Upsilon, \Upsilon \in \Gamma$, 故$\Omega \in \bigcap\limits_{\Upsilon \in \Gamma} \mathscr{A}_\Upsilon = \mathscr{A}$; 其次, 若 $A \in \mathscr{A}$, 则对一切 $\Upsilon \in \Gamma, A \in \mathscr{A}_\Upsilon$, 由 $\mathscr{A}_\Upsilon$ 是 σ-代数知 $A^c \in \mathscr{A}_\Upsilon, \Upsilon \in \Gamma$, 因而 $A^c \in \bigcap\limits_{\Upsilon \in \Gamma} \mathscr{A}_\Upsilon$; 第三, 若 $A_n \in \mathscr{A}, n = 1, 2, \cdots$, 则对每一 $\Upsilon \in \Gamma$ 来说, $A_n \in \mathscr{A}_\Upsilon, n = 1, 2, \cdots$, 故由 $\mathscr{A}_\Upsilon$ 是 σ-代数知 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{A}_\Upsilon, \Upsilon \in \Gamma$, 因而 $\bigcup\limits_{n=1}^{\infty} A_n \in \bigcap\limits_{\Upsilon \in \Gamma} \mathscr{A}_\Upsilon = \mathscr{A}$, 总结以上三点, 即知 $\mathscr{A}$ 为一 σ-代数. □

我们现在来证明, Ω 的任一子集类总可以扩张成为 Ω 中的一个 σ-代数, 这就是

定理 2　设 $\mathfrak{C}$ 是 Ω 的一个子集类, 则有一 Ω 中的 σ-代数 $\mathscr{A}_0$ 存在, 它具有下列性质:

i) $\mathfrak{C} \subset \mathscr{A}_0$;

ii) 若 $\mathscr{A}$ 为 Ω 中的 σ-代数且 $\mathfrak{C} \subset \mathscr{A}$, 则 $\mathscr{A}_0 \subset \mathscr{A}$,

并且具有性质 i), ii) 的 σ-代数由 $\mathfrak{C}$ 唯一决定.

证 由于 Ω 的一切子集构成的集类是一 σ-代数且显然包含 $\mathfrak{E}$, 故 Ω 中包含 $\mathfrak{E}$ 的 σ-代数的类是不空的, 令 $\mathscr{A}_0$ 是 Ω 中包含 $\mathfrak{E}$ 的一切 σ-代数的交, 则由性质 7 知 $\mathscr{A}_0$ 是 Ω 中的一个 σ-代数且显然包含 $\mathfrak{E}$, 即$\mathscr{A}_0$ 满足 i), 而由于 $\mathscr{A}_0$ 的定义知它具有性质 ii).

其次, 设 $\mathscr{A}_1$ 是具有性质 i), ii) 的任一 σ-代数, 则因 $\mathscr{A}_1$ 是包含 $\mathfrak{E}$ 的 σ-代数, 由 ii) 知 $\mathscr{A}_0\subset\mathscr{A}_1$, 另一方面由于 $\mathscr{A}_0$ 是包含 $\mathfrak{E}$ 的 σ-代数, 而 $\mathscr{A}_1$ 具有性质 ii), 故 $\mathscr{A}_1\subset\mathscr{A}_0$, 因此 $\mathscr{A}_1=\mathscr{A}_0$, 即具有性质 i), ii) 的 σ-代数由 $\mathfrak{E}$ 唯一决定. □

满足定理 2 中性质 i), ii) 的 σ-代数在今后的讨论中具有重要的地位, 因而我们给出下面的

定义 4 设 $\mathfrak{E}$ 是空间 Ω 中的一个集类, 则具有定理 2 中性质 i), ii) 的 Ω 中的 σ-代数称为 **$\mathfrak{E}$上的最小 σ-代数**也称为由 $\mathfrak{E}$ 生成的 σ-代数, 记作 $\sigma(\mathfrak{E})$.

容易证明, 半集代数 $\mathscr{S}$ 上的最小 σ-代数 $\sigma(\mathscr{S})$ 与由定理 1 中所定义的 $\mathscr{S}$ 上最小集代数 $\mathscr{F}$ 上的最小 σ-代数 $\sigma(\mathscr{F})$ 相等, 即

定理 3 若 $\mathscr{S}$ 是 Ω 中的半集代数, $\mathscr{F}$ 是按定理 1 中定义的 $\mathscr{S}$ 上的最小集代数, 则

$$\sigma(\mathscr{S})=\sigma(\mathscr{F}).$$

证 由定理 1 知

$$\mathscr{F}=\left\{A:A=\sum_{k=1}^{n}A_k,A_k\in\mathscr{S},k=1,\cdots,n\text{两两不交},n\geqslant 1\right\}.$$

故 $\mathscr{S}\subset\mathscr{F}\subset\sigma(\mathscr{F})$, 再由 $\sigma(\mathscr{F})$ 是含 $\mathscr{S}$ 的 σ-代数及 $\sigma(\mathscr{S})$ 的最小性知 $\sigma(\mathscr{S})\subset\sigma(\mathscr{F})$. 而 $\sigma(\mathscr{S})$ 是 σ-代数, 故必为集代数且包含 $\mathscr{S}$, 故 $\sigma(\mathscr{S})\supset\mathscr{F}$, 而 $\sigma(\mathscr{F})$ 是包含 $\mathscr{F}$ 的最小 σ-代数, 故 $\sigma(\mathscr{S})\supset\sigma(\mathscr{F})$, 从而有

$$\sigma(\mathscr{F})=\sigma(\mathscr{S}).$$ □

我们举出两个重要的 σ-代数的例子.

例 7 (定义) 例 3 中给出的在实直线上的区间类作成的半集代数 $\mathscr{S}^{(1)}$, 其上最小 σ-代数 $\sigma(\mathscr{S}^{(1)})$ 称为**直线上的 Borel 域**, 记作 $\mathscr{B}^{(1)}$ 或 $\mathscr{B}$.$\mathscr{B}^{(1)}$ 中的元称为**直线上的 Borel 集**;例 4 中的平面上的矩形类作成的半集代数 $\mathscr{S}^{(2)}$, 其上最小 σ-代数 $\sigma(\mathscr{S}^{(2)})$ 称为**平面上的 Borel 域**, 记作 $\mathscr{B}^{(2)}$,$\mathscr{B}^{(2)}$ 中的元称为**平面上的 Borel 集或二维 Borel 集**.类似地可以定义n **维 Borel 域**, 记作 $\mathscr{B}^{(n)}$,$\mathscr{B}^{(n)}$ 中的元称为n **维 Borel 集**.

又令 $\mathfrak{E}_1'=\{(-\infty,a),a\in R\}$, 则 $\sigma(\mathfrak{E}_1')=\sigma(\mathscr{S}^{(1)})$; $\mathfrak{E}_2'=\{(-\infty,a):a\in R^{(2)}\}$, 其中 $a=(a_1,a_2),(-\infty,a)=\{(x,y):-\infty<x<a_1,-\infty<y<a_2\}$, 则 $\sigma(\mathfrak{E}_2')=\sigma(\mathscr{S}^{(2)})$, 对 n 维 Borel 域有类似的结论.

例 8 (定义)　设 $\Omega = Z^{(1)}$ 是复平面, 即

$$Z^{(1)} = \{x + iy : x, y \in R\}.$$

令

$$\mathfrak{E} = \{A : A = \{x + iy : a_1 \leqslant x < b_1, a_2 \leqslant y < b_2\},$$
$$a_1, a_2, b_1, b_2 \in R\},$$

称 $\sigma(\mathfrak{E})$ 为**复 Borel 域**, 记作 $\mathscr{B}_z^{(1)}$. 事实上一个复 Borel 域和一个二维实 Borel 域是同构的. 同样还可以定义n **维复 Borel 域**, 记作 $\mathscr{B}_z^{(n)}$. 一个 n 维复 Borel 域和一个 $2n$ 维实 Borel 域也是同构的.

例 9 (定义)　设 $\Omega = \widetilde{R}^{(n)}$ 是广义 n 维实空间 (包含一切形如 $(x_1, \cdots, x_n), x_i = +\infty$ 或 $-\infty$ 或任意实数 $i = 1, \cdots, n$ 的点), 若令

$$\widetilde{\mathfrak{E}}_n = \{[a, b] : a, b \in \widetilde{R}^{(n)}\}$$

$[a, b] = \{(x_1, \cdots, x_n) : a_i \leqslant x_i \leqslant b_i, i = 1, \cdots, n\}$, 则以 $\widetilde{R}^{(n)}$ 为空间 $\mathfrak{E}_n$ 上的最小 σ-代数称为**广义 n 维 Borel 域**, 记作 $\widetilde{\mathscr{B}}^{(n)}, \widetilde{\mathscr{B}}^{(n)}$ 中的元称为**广义 n 维 Borel 集**.类似地可以定义**广义 n 维复 Borel 域**, 记作 $\widetilde{\mathscr{B}}_z^{(n)}$. 和**广义 n 维复 Borel 集.**

1.3.3　单调类定理

在实际问题中, 检验一个集类是否是 σ-代数往往比较困难, 我们把 σ-代数与以下定义的单调类、π-系、λ-系联系起来, 问题就比较容易解决, 于是引入

定义 5　设 $\mathfrak{M}$ 是 Ω 中的子集构成的集类, 具有性质:

I) 若 $A_n \in \mathfrak{M}, n = 1, 2, \cdots$, 且 $A_1 \subset A_2 \subset \cdots$, 则 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathfrak{M}$;

II) 若 $A_n \in \mathfrak{M}, n = 1, 2, \cdots$, 且 $A_1 \supset A_2 \supset \cdots$, 则 $\bigcap\limits_{n=1}^{\infty} A_n \in \mathfrak{M}$, 则称 $\mathfrak{M}$ 为 Ω 中的一个**单调类.**

今后用 $A_n \uparrow$ 表示集序列 $A_n, n = 1, 2, \cdots$ 是不降的, 即 $A_1 \subset A_2 \subset \cdots$, 用 $A_n \downarrow$ 表示 $A_n, n = 1, 2, \cdots$ 是不增的, 即 $A_1 \supset A_2 \supset \cdots$.

例 10　设 $\Omega = R^{(1)}, \mathfrak{M} = \{(-\infty, a), (-\infty, a], (-\infty, +\infty), \varnothing : a \in R^{(1)}\}$, 则 $\mathfrak{M}$ 是单调类. 但它对减法运算不封闭. 例如 $A = (-\infty, a], B = (-\infty, b), b > a$, 则 $B - A = (a, b) \notin \mathfrak{M}$, 因而 $\mathfrak{M}$ 不是 σ-代数.

从例 10 可知, Ω 中的单调类并不一定是 σ-代数, 但是 Ω 中的任一 σ-代数都是单调类, 这是因为 σ-代数对可数并和可数交是封闭的. 所以我们说单调类受限制比 σ-代数少, 虽然如此, 单调类与 σ-代数之间有着下列定理所陈述的关系.

定理 4　Ω 中的任一 σ-代数是单调类, 若 $\mathscr{F}$ 是 Ω 中的集代数且为单调类, 则它是 σ-代数.

证 由单调类的定义易见定理的前半部是正确的. 今证定理的后半部, 设 $\mathscr{F}$ 是 Ω 中的集代数, 并且它又是单调类, 于是要证 $\mathscr{F}$ 是 σ-代数, 只需要证明它对可数并是封闭的即可.

设 $A_n \in \mathscr{F}, n=1,2,\cdots$, 则令 $B_n = \bigcup\limits_{k=1}^{n} A_k$, 由于 $\mathscr{F}$ 是集代数, 故 $B_n \in \mathscr{F}, n = 1,2,\cdots$, 且 $B_n \uparrow$, 又由于 $\mathscr{F}$ 是单调类, 因而 $\bigcup\limits_{n=1}^{\infty} B_n \in \mathscr{F}$, 但 $\bigcup\limits_{n=1}^{\infty} A_n = \bigcup\limits_{n=1}^{\infty} B_n$, 故 $\mathscr{F}$ 对可数并封闭, 因此 $\mathscr{F}$ 是 σ-代数.

和对 σ-代数的讨论一样, 容易证明单调类具有

性质 8 Ω 中的一切子集构成的集类是一单调类;

性质 9 Ω 中任意多个单调类的交仍然是 Ω 中的单调类.

从而有

定理 5 设 $\mathfrak{E}$ 是 Ω 中任一集类, 则有一 Ω 中的单调类 $\mathfrak{M}_0$ 存在, 它具有下列性质:

i) $\mathfrak{E} \subset \mathfrak{M}_0$;

ii) 若 $\mathfrak{M}$ 为 Ω 中的单调类, 且 $\mathfrak{E} \subset \mathfrak{M}$, 则 $\mathfrak{M}_0 \subset \mathfrak{M}$,

并且具有性质 i), ii) 的单调类是由 $\mathfrak{E}$ 唯一决定的, 这个单调类我们称之为 $\mathfrak{E}$ 上的**最小单调类**, 记作 $\mathfrak{M}(\mathfrak{E})$.

关于集代数上的最小单调类与最小 σ-代数之间的关系, 有重要的

定理 6 设 $\mathscr{F}$ 是 Ω 中的一个集代数, 则 $\mathfrak{M}(\mathscr{F}) = \sigma(\mathscr{F})$, 因此包含 $\mathscr{F}$ 的任一单调类必包含 $\sigma(\mathscr{F})$.

证 由于 $\sigma(\mathscr{F}) \supset \mathscr{F}$, 且由定理 4 知 $\sigma(\mathscr{F})$ 是一单调类, 因此

$$\sigma(\mathscr{F}) \supset \mathfrak{M}(\mathscr{F}).$$

如果我们能够证明 $\mathfrak{M}(\mathscr{F})$ 是一 σ-代数, 则由 $\mathfrak{M}(\mathscr{F}) \supset \mathscr{F}$ 及 $\sigma(\mathscr{F})$是$\mathscr{F}$ 上的最小 σ-代数知 $\sigma(\mathscr{F}) \subset \mathfrak{M}(\mathscr{F})$, 定理即获证明.

为要证明 $\mathfrak{M}(\mathscr{F})$ 是 σ-代数, 由定理 4 知只需证明 $\mathfrak{M}(\mathscr{F})$ 是一集代数即可. 由于 $\Omega \in \mathscr{F} \subset \mathfrak{M}(\mathscr{F})$, 故由性质 3 知, 要证 $\mathfrak{M}(\mathscr{F})$ 为一集代数, 只需证明对任何 $A, B \in \mathfrak{M}(\mathscr{F})$, 都有 $A - B \in \mathfrak{M}(\mathscr{F})$. 以下往证此事, 这一证明的手法希望读者加以注意.

对于任一给定的 $A \in \mathfrak{M}(\mathscr{F})$, 令

$$\mathfrak{M}_A \triangleq \{B : B \in \mathfrak{M}(\mathscr{F}), A - B, B - A \in \mathfrak{M}(\mathscr{F})\}. \tag{1}$$

显然若能证明对于每一 $A \in \mathfrak{M}(\mathscr{F})$, 都有 $\mathfrak{M}_A = \mathfrak{M}(\mathscr{F})$, 则我们的证明即已完成. 为了证明对每一 $A \in \mathfrak{M}(\mathscr{F})$ 都有 $\mathfrak{M}_A = \mathfrak{M}(\mathscr{F})$, 我们依次来证明下列诸事实: (a)

对每一 $A \in \mathfrak{M}(\mathscr{F}), \mathfrak{M}_A$ 是一单调类;(b) 若 $A \in \mathscr{F}$, 则 $\mathfrak{M}_A = \mathfrak{M}(\mathscr{F})$;(c) 对任何 $A \in \mathfrak{M}(\mathscr{F}), \mathfrak{M}_A = \mathfrak{M}(\mathscr{F})$.

(a) 设 A 是取定的集合, 若 $\{B_n\} \subset \mathfrak{M}_A$ 且 $B_n \uparrow$,则由$\mathfrak{M}_A$ 的定义知 $A - B_n, B_n - A, B_n \in \mathfrak{M}(\mathscr{F})$, 且 $B_n \uparrow, A - B_n \downarrow, B_n - A \uparrow$, 由于$\mathfrak{M}(\mathscr{F})$ 是单调类, 故

$$\bigcup_{n=1}^{\infty} B_n \in \mathfrak{M}(\mathscr{F});$$

$$\begin{aligned} A - \bigcup_{n=1}^{\infty} B_n &= A\left(\bigcap_{n=1}^{\infty} B_n^c\right) = \bigcap_{n=1}^{\infty}(A \cap B_n^c) \\ &= \bigcap_{n=1}^{\infty}(A - B_n) \in \mathfrak{M}(\mathscr{F}); \end{aligned}$$

$$\left(\bigcup_{n=1}^{\infty} B_n\right) - A = \bigcup_{n=1}^{\infty}(B_n - A) \in \mathfrak{M}(\mathscr{F}).$$

由 $\mathfrak{M}_A$ 的定义知 $\bigcup\limits_{n=1}^{\infty} B_n \in \mathfrak{M}_A$.

同理可证, 若 $\{B_n\} \subset \mathfrak{M}_A, B_n \downarrow$, 则 $\bigcap\limits_{n=1}^{\infty} B_n \in \mathfrak{M}_A$. 因此 $\mathfrak{M}_A$ 是一单调类.

(b) 若 $A \in \mathscr{F}$, 则对任何 $B \in \mathscr{F}$,有$B - A \in \mathscr{F}, A - B \in \mathscr{F}$, 但 $\mathscr{F} \subset \mathfrak{M}(\mathscr{F})$, 于是$B, A - B, B - A \in \mathfrak{M}(\mathscr{F})$, 故由$\mathfrak{M}_A$ 的定义知 $B \in \mathfrak{M}_A$, 因此 $\mathfrak{M}_A \supset \mathscr{F}$, 由 (a) 知 $\mathfrak{M}_A$ 是包含 $\mathscr{F}$ 的单调类, 故 $\mathfrak{M}(\mathscr{F}) \subset \mathfrak{M}_A$, 但由 (1) 知 $\mathfrak{M}_A \subset \mathfrak{M}(\mathscr{F})$, 故

$$\mathfrak{M}_A = \mathfrak{M}(\mathscr{F}), \text{对任一} A \in \mathscr{F}.$$

(c) 若 $B \in \mathfrak{M}(\mathscr{F})$,则对任一$A \in \mathscr{F}$ 来说, 由 (b) 知 $\mathfrak{M}_A = \mathfrak{M}(\mathscr{F})$,因而$B \in \mathfrak{M}_A$, 由 (1) 知$A - B, B - A \in \mathfrak{M}(\mathscr{F})$,又由$\mathscr{F} \subset \mathfrak{M}(\mathscr{F})$知$A \in \mathfrak{M}(\mathscr{F})$, 故再由 (1) 知 $A \in \mathfrak{M}_B$, 这就是说, 对于任何 $A \in \mathscr{F}$ 来说, 都有 $A \in \mathfrak{M}_B$, 亦即 $\mathscr{F} \subset \mathfrak{M}_B$, 由 (a) 知 $\mathfrak{M}_B$ 是单调类, 故 $\mathfrak{M}(\mathscr{F}) \subset \mathfrak{M}_B$, 因而由 $\mathfrak{M}_B \subset \mathfrak{M}(\mathscr{F})$ 知

$$\mathfrak{M}_B = \mathfrak{M}(\mathscr{F}), \text{对任一} B \in \mathfrak{M}(\mathscr{F}).$$

由上面的分析知定理证毕. □

这个定理告诉我们要研究集代数 $\mathscr{F}$ 上的最小 σ-代数, 只要研究 $\mathscr{F}$ 上的最小单调类就行了. 由于单调类所受限制比 σ-代数要少, 因此在许多场合下, 是比较容易研究的.

在有些场合, 验证某集类包含某集代数的单调类还比较困难或麻烦, 需要引入下面的

定义 6 若 Ω 的子集类 $\mathfrak{E}$ 具有性质: 若 $A, B \in \mathfrak{E}$, 则 $A \cap B \in \mathfrak{E}$, 则称 $\mathfrak{E}$ 为 π-系; 若 $\mathfrak{E}$ 满足以下条件:

I) $\Omega \in \mathfrak{E}$;

II) 若 $A, B \in \mathfrak{E}, A \subset B$, 则 $B - A \in \mathfrak{E}$;(对真差封闭)

III) 若 $A_n \in \mathfrak{E}, n = 1, 2 \cdots$ 且 $A_n \uparrow$, 则 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathfrak{E}$, (对不降序列的并封闭),

则称 $\mathfrak{E}$为$\lambda-$系.

性质 10 若 $\mathfrak{E}$ 是 λ-系, 则 $\mathfrak{E}$ 是单调类.

证 只要证 $\mathfrak{E}$ 对不增序列的交封闭即可. 若 $A_n \downarrow$, 则由 I), II) 知 $A_n^c \in \mathfrak{E}, n = 1, 2 \cdots$ 且 $A_n^c \uparrow$, 由 III) 知 $\bigcup\limits_{n=1}^{\infty} A_n^c \in \mathfrak{E}$, 再由 I), II) 知 $\bigcap\limits_{n=1}^{\infty} A_n = \left(\bigcup\limits_{n=1}^{\infty} A_n^c\right)^c \in \mathfrak{E}$. □

性质 11 若 $\mathfrak{E}$ 同时为 π-系和 λ-系, 则 $\mathfrak{E}$为σ- 代数.

证 由性质 10 知若 $\mathfrak{E}$ 为 λ-系则必为单调类, 因此只需证 $\mathfrak{E}$ 是集代数, 由 λ-系的定义 I), II) 及性质 3 知, 只需证 $\mathfrak{E}$ 对交封闭, 而 $\mathfrak{E}$ 为 π-系因而对交封闭.

和最小 σ-代数、最小单调类的存在性一样, 可以证明对任何 Ω 的子集类 $\mathfrak{E}$, 存在 $\mathfrak{E}$ 上的最小 λ-系 λ_0, 满足

i) $\mathfrak{E} \subset \lambda_0$;

ii) 若 λ 是 Ω 上包含 $\mathfrak{E}$ 的 λ-系, 则 $\lambda \supset \lambda_0$. 且满足 i), ii) 的 λ-系由 $\mathfrak{E}$ 唯一决定.

π-系上的 λ-系与 π-系上的最小 σ-代数有下述关系:

定理 7 设 $\mathfrak{E}$ 是 π-系, $\lambda(\mathfrak{E})$ 是 $\mathfrak{E}$ 上最小 λ-系, 则

$$\lambda(\mathfrak{E}) = \sigma(\mathfrak{E}).$$

证 Ω 中的任一 σ-代数必为 λ-系, 故 $\sigma(\mathfrak{E})$ 是包含 $\mathfrak{E}$ 的 λ-系, 因而 $\sigma(\mathfrak{E}) \supset \lambda(\mathfrak{E})$; 再由性质 11 知为要证明 $\lambda(\mathfrak{E}) \supset \sigma(\mathfrak{E})$, 只要证明 $\lambda(\mathfrak{E})$ 是 π-系即可. 令

$$\mathscr{A}_A \triangleq \{B : B \in \lambda(\mathfrak{E}),\ \text{且}\ A \cap B \in \lambda(\mathfrak{E})\}. \tag{2}$$

为要证明 $\lambda(\mathfrak{E})$ 是 π-系, 只要证明对任何 $A \in \lambda(\mathfrak{E}), \mathscr{A}_A = \lambda(\mathfrak{E})$. 证明的手法和定理 6 一样, 分三个步骤:

(a) 证明对一切 $A \in \lambda(\mathfrak{E}), \mathscr{A}_A$ 是 λ-系: 事实上, 因为 $A \cap \Omega = A \in \lambda(\mathfrak{E})$, 故 $\Omega \in \mathscr{A}_A$; 若 $B, C \in \mathscr{A}_A$ 且 $B \subset C$, 由 $\mathscr{A}_A$的定义知$A \cap B, A \cap C \in \lambda(\mathfrak{E})$ 且 $A \cap B \subset A \cap C$, 由 $\lambda(\mathfrak{E})$ 是 λ-系知 $A \cap (C - B) = (A \cap C) - (A \cap B) \in \lambda(\mathfrak{E})$, 故 $C - B \in \mathscr{A}_A$, 即 $\mathscr{A}_A$ 对真差封闭; 再若 $B_n \in \mathscr{A}_A, n = 1, 2, \cdots B_n \uparrow$, 则 $A \cap B_n \in \lambda(\mathfrak{E}), n = 1, 2, \cdots$ 且 $(A \cap B_n) \uparrow$, 故由 $\lambda(\mathfrak{E})$ 是 λ-系知 $A \cap \left(\bigcup\limits_{n=1}^{\infty} B_n\right) = \bigcup\limits_{n=1}^{\infty} (A \cap B_n) \in \lambda(\mathfrak{E})$, 故对一切 $A \in \lambda(\mathfrak{E}), \mathscr{A}_A$是$\lambda$-系.

(b) 证明若 $A \in \mathfrak{E}$, 则 $\mathscr{A}_A = \lambda(\mathfrak{E})$: 若$B \in \mathfrak{E}$,则由 $\mathfrak{E}$ 是 π-系知 $A \cap B \in \mathfrak{E} \subset \lambda(\mathfrak{E})$.故$\mathfrak{E} \subset \mathscr{A}_A$, 由于 $\mathscr{A}_A$, 是 λ-系, 故 $\lambda(\mathfrak{E}) \subset \mathscr{A}_A$, 再由 $\mathscr{A}_A$ 的定义知 $\lambda(\mathfrak{E}) = \mathscr{A}_A$.

(c) 证明若 $A \in \lambda(\mathfrak{E})$, 则 $\mathscr{A}_A = \lambda(\mathfrak{E})$: 若 $B \in \mathfrak{E}$, 则由 (b) 知 $A \in \mathscr{A}_B$, 由于对称性, $B \in \mathscr{A}_A$, 故对任何 $B \in \mathfrak{E}$ 都有 $B \in \mathscr{A}_A$, 亦即 $\mathfrak{E} \subset \mathscr{A}_A$, 由 (a) 知 $\mathscr{A}_A$ 是 λ-系, 故 $\lambda(\mathfrak{E}) \subset \mathscr{A}_A$, 因而由 $\mathscr{A}_A \subset \lambda(\mathfrak{E})$ 知 $\mathscr{A}_A = \lambda(\mathfrak{E})$.

根据前面的分析, 定理证毕. □

下一定理是定理 6, 7 的综合, 有时也称为**集合形式的单调类定理**, 它在验证一个集类是 σ-代数时经常用到.

定理 8　设 $\mathscr{C}, \mathscr{A}$为Ω的两个子集类且$\mathscr{C} \subset \mathscr{A}$.

i) 若 $\mathscr{A}$为λ-系,$\mathscr{C}$为π-系, 则 $\sigma(\mathscr{C}) \subset \mathscr{A}$.

ii) 若 $\mathscr{A}$ 为单调类, $\mathscr{C}$ 为集代数, 则 $\sigma(\mathscr{C}) \subset \mathscr{A}$.

证　i) 由 $\mathscr{A} \supset \mathscr{C}$ 且 $\mathscr{A}$ 是 λ-系知 $\mathscr{A} \supset \lambda(\mathscr{C})$,再若$\mathscr{C}$是$\pi$-系,由定理 7 知$\lambda(\mathscr{C}) = \sigma(\mathscr{C})$,故得$\mathscr{A} \supset \sigma(\mathscr{C})$.

ii) 由 $\mathscr{A} \supset \mathscr{C}$ 及 $\mathscr{A}$ 为单调类知 $\mathscr{A} \supset \mathfrak{M}(\mathscr{C})$, 再若 $\mathscr{C}$ 是集代数, 由定理 6 知 $\mathfrak{M}(\mathscr{C}) = \sigma(\mathscr{C})$, 因此 $\mathscr{A} \supset \sigma(\mathscr{C})$.

这个定理用法如下: 通常已知 $\mathscr{C}$ 中元具有性质 S, 需要证明 $\sigma(\mathscr{C})$ 中元亦具有性质 S, 为此令

$$\Lambda = \{B : B\text{具有性质}S\},$$

然后证明 $\mathscr{C}$ 是 π-系 (或集代数), 再证 Λ 是 λ-系 (相应地: 单调类), 于是 $\Lambda \supset \sigma(\mathscr{C})$, 即证明了 $\sigma(\mathscr{C})$ 中的元都具有性质 S. 这种方法称为 λ-系 (相应地: 单调类) 方法.

1.3.4　乘积空间与乘积 σ-代数

现在我们引入一种不同类型集合间的运算以及相应的若干概念. 设 Ω_1, Ω_2 是两个给定的空间 (集合), 以下如无特别声明, 用 ω_1 和 ω_2 分别表示 Ω_1 和 Ω_2 的元素, $A_1, B_1, \cdots$ 和 $A_2, B_2, \cdots$ 分别表示 Ω_1 和 Ω_2 的子集.

定义 7　集 A_1 和 A_2 的**乘积集**, 记作 $A_1 \times A_2$, 指的是下面的集合:

$$A_1 \times A_2 = \{(\omega_1, \omega_2) : \omega_1 \in A_1, \omega_2 \in A_2\}.$$

称 $\Omega_1 \times \Omega_2$ 为乘积空间. 对任意的 $A_i \in \Omega_i, i = 1, 2, A_1 \times A_2$ 都是 $\Omega_1 \times \Omega_2$ 的子集.$A_1 \times A_2$ 也称为 $\Omega_1 \times \Omega_2$ 中的**矩形**.

性质 12　乘积集具有下述性质:

(i) $A_1 \times A_2 = \varnothing \Longleftrightarrow A_1 = \varnothing$或$A_2 = \varnothing$;

(ii) 若 $A_1 \times A_2$ 非空, 则

$$A_1 \times A_2 \subset B_1 \times B_2 \Longleftrightarrow A_1 \subset B_1 \text{ 且 } A_2 \subset B_2,$$

$$A_1 \times A_2 = B_1 \times B_2 \Longleftrightarrow A_1 = B_1 \text{ 且 } A_2 = B_2;$$

(iii) $(A_1 \times A_2) \cap (B_1 \times B_2) = (A_1 \cap B_1) \times (A_2 \cap B_2)$;

(iv) $(A_1 \times A_2)^c = (A_1^c \times A_2) + (A_1 \times A_2^c) + (A_1^c \times A_2^c)$,其中$(A_1 \times A_2)^c = (\Omega_1 \times \Omega_2) - (A_1 \times A_2), A_1^c = \Omega_1 - A_1, A_2^c = \Omega_2 - A_2$.

注意 (iv) 中等式右端各矩形互不相交.

定理 9 设 $\mathscr{S}_i$ 是 Ω_i 中的半集代数, $i = 1, 2$, 则

i) $\mathfrak{C} \hat{=} \{A_1 \times A_2 : A_i \in \mathscr{S}_i, i = 1, 2\}$ 是 $\Omega_1 \times \Omega_2$ 中的半集代数;

ii) $\mathfrak{C}' \hat{=} \left\{ \sum_{i=1}^{n} A^{(i)} : A^{(i)} \in \mathfrak{C}, i = 1, \cdots, n \text{两两不交}, n \geqslant 1 \right\}$ 是 $\Omega_1 \times \Omega_2$ 中的集代数.

证 i) 因为 $\Omega_i \in \mathscr{S}_i, i = 1, 2$ 由 $\mathfrak{C}$ 的定义知 $\Omega_1 \times \Omega_2 \in \mathfrak{C}$, 再由性质 12 的 (iii) 知 $\mathfrak{C}$ 对交封闭, 由 (iv) 及 $A_i^c, i = 1, 2$ 可表成 $\mathscr{S}_i$ 中不交集的和可知 $\mathfrak{C}$ 中集的余集可表成 $\mathfrak{C}$ 中有限个不交集的和, 因而 $\mathfrak{C}$ 是半集代数.

ii) 由定理 1 立刻得到 $\mathfrak{C}'$ 是集代数. □

定义 8 设 $\mathscr{A}_i$ 是 Ω_i 中的 σ-代数, $i = 1, 2$, 称包含集类

$$\mathfrak{C} \hat{=} \{A_1 \times A_2 : A_i \in \mathscr{A}_i, i = 1, 2\}$$

的最小 σ-代数 $\sigma(\mathfrak{C})$ 为$\mathscr{A}_1$ **和** $\mathscr{A}_2$ **的乘积** σ**-代数**, 记作 $\mathscr{A}_1 \times \mathscr{A}_2$.

为了叙述方便, 称 $\mathfrak{C}$ 中的元为可测矩形.

这些概念当 $\Omega_1 = \Omega_2 = R^{(1)}$ 时, 都是我们所熟悉的.

例 11 设 $\Omega_1 = \Omega_2 = R^{(1)}, \mathscr{A}_1 = \mathscr{A}_2 = \mathscr{B}^{(1)}$(直线上的 Borel 域), 则二维实空间 $R^{(2)} = R^{(1)} \times R^{(1)}$, 而平面上的二维 Borel 域 $\mathscr{B}^{(2)} = \mathscr{B}^{(1)} \times \mathscr{B}^{(1)}$.

事实上, 由例 7 知, $\mathscr{B}^{(1)} \times \mathscr{B}^{(1)}$, 包含一切下述形式的矩形:

$$\{(x_1, x_2) \in R^{(2)} : -\infty < x_1 < a_1, -\infty < x_2 < a_2\} = (-\infty, a_1) \times (-\infty, a_2),$$

而 $\mathscr{B}^{(2)}$ 是包含一切上述矩形类的最小 σ-代数, 故

$$\mathscr{B}^{(2)} \subset \mathscr{B}^{(1)} \times \mathscr{B}^{(1)}.$$

另一方面, 对任何 $a_2 \in R^{(1)}$, 令

$$\mathfrak{M}_1 = \{A_1 : A_1 \times (-\infty, a_2) \in \mathscr{B}^{(2)}\}.$$

显然 $\mathfrak{M}_1$ 包含一切区间 $(-\infty, a_1)$, 并且 $\mathfrak{M}_1$ 包含 $R^{(1)}$; 同时若 $A_1^{(1)}, A_1^{(2)}, \cdots \in \mathfrak{M}_1$, 则

$$A_1^{(m)} \times (-\infty, a_2) \in \mathscr{B}^{(2)}, m = 1, 2, \cdots,$$

因而
$$\left(\bigcup_{m=1}^{\infty} A_1^{(m)}\right) \times (-\infty, a_2) = \bigcup_{m=1}^{\infty} (A_1^{(m)} \times (-\infty, a_2)) \in \mathscr{B}^{(2)}, \text{ 即 } \bigcup_{m=1}^{\infty} A_1^{(m)} \in \mathfrak{M}_1;$$
又若 $A_1 \in \mathfrak{M}_1$, 则
$$A_1 \times (-\infty, a_2) \in \mathscr{B}^{(2)}.$$
因而
$$A_1^c \times (-\infty, a_2) = [(-\infty, \infty) \times (-\infty, a_2)] - [A_1 \times (-\infty, a_2)] \in \mathscr{B}^{(2)},$$
即 $A_1^c \in \mathfrak{M}_1$. 这样 $\mathfrak{M}_1$ 是包含一切区间 $(-\infty, a_1)$ 的 σ-代数, 因而 $\mathfrak{M}_1 \supset \mathscr{B}^{(1)}$, 由此知 $\mathscr{B}^{(2)}$ 包含一切形如
$$A_1 \times (-\infty, a_2), \quad A_1 \in \mathscr{B}^{(1)}, a_2 \in R^{(1)}$$
的矩形作成的类.

再令
$$\mathfrak{M}_2 = \{A_2 : A_1 \times A_2 \in \mathscr{B}^{(2)}, A_1 \in \mathscr{B}^{(1)}\},$$
类似地可以证明 $\mathfrak{M}_2 \supset \mathscr{B}^{(1)}$, 从而又有 $\mathscr{B}^{(2)}$ 包含一切形如
$$A_1 \times A_2, A_i \in \mathscr{B}^{(1)}, i = 1, 2$$
的矩形作成的类, 故
$$\mathscr{B}^{(2)} \supset \mathscr{B}^{(1)} \times \mathscr{B}^{(1)}.$$
于是得出
$$\mathscr{B}^{(2)} = \mathscr{B}^{(1)} \times \mathscr{B}^{(1)}.$$

对于复 Borel 域及广义实、复 Borel 域有类似的结论: $\mathscr{B}_z^{(2)} = \mathscr{B}_z^{(1)} \times \mathscr{B}_z^{(1)}, \widetilde{\mathscr{B}}^{(2)} = \widetilde{\mathscr{B}}^{(1)} \times \widetilde{\mathscr{B}}^{(1)}$ 及 $\widetilde{\mathscr{B}}_z^{(2)} = \widetilde{\mathscr{B}}_z^{(1)} \times \widetilde{\mathscr{B}}_z^{(1)}$.

下面我们叙述有限维乘积空间、乘积集及乘积 σ-代数的定义.

设 $\Omega_i, i = 1, \cdots, n$是$n$个空间 (集合), 以下如无特别声明, 我们用 ω_i 表示 Ω_i 的元素, A_i 表示 Ω_i 的子集, $i = 1, \cdots, n$. 所谓集 $A_1, \cdots, A_n$ 的乘积集, 记作 $A_1 \times \cdots \times A_n$或 $\prod_{i=1}^{n} A_i$, 指的是下面的集合:
$$\prod_{i=1}^{n} A_i = A_1 \times \cdots \times A_n = \{(\omega_1, \cdots, \omega_n) : \omega_i \in A_i, i = 1, \cdots, n\}.$$
显然, 乘积集 $A_1 \times \cdots \times A_n$ 是空间 $\Omega_1 \times \cdots \times \Omega_n$ 的子集, 也称为 $\Omega_1 \times \cdots \times \Omega_n$ 中的矩形, 如果 $\mathscr{A}_i$ 是 Ω_i 中的 σ-代数, $i = 1, \cdots, n$, 则
$$A_1 \times \cdots \times A_n, A_i \in \mathscr{A}_i, i = 1, \cdots, n$$

称为**可测矩形**. 可测矩形类为半集代数. 由一切有限个不相交的可测矩形的和所成的类 $\mathscr{F}$ 是一个 $\Omega_1\times\cdots\times\Omega_n$ 中的集代数, 包含一切可测矩形的最小 σ-代数称为 $\mathscr{A}_1,\cdots,\mathscr{A}_n$ 的乘积 σ-代数, 记作 $\mathscr{A}_1\times\cdots\times\mathscr{A}_n$或 $\prod\limits_{i=1}^{n}\mathscr{A}_i$.

关于乘积集的运算并不满足结合律, 例如

$$\begin{aligned}A_1\times A_2\times A_3&=\{(\omega_1,\omega_2,\omega_3,):\omega_i\in A_i,i=1,2,3\};\\(A_1\times A_2)\times A_3&=\{((\omega_1,\omega_2),\omega_3):(\omega_1\times\omega_2)\in A_1\times A_2,\omega_3\in A_3\}\\&=\{((\omega_1,\omega_2),\omega_3):\omega_i\in A_i,i=1,2,3\}.\end{aligned}$$

显然它们是由不同的元素组成的. 但是从它们元素的结构来看, 都是由 A_1 中的 ω_1,A_2 中的 ω_2 及 A_3 中的 ω_3 构成的, 而且它们出现的先后顺序相同. 只是构成的手续有所不同罢了. 对于我们所讨论的问题来说, 重要的是元素的基本结构, 而无需对于构成的手续严加区别. 因此我们将 $(\omega_1,\omega_2,\omega_3)$ 和 $((\omega_1,\omega_2),\omega_3)$ 看作是一样的, 因而认为

$$\Omega_1\times\Omega_2\times\Omega_3=(\Omega_1\times\Omega_2)\times\Omega_3=\Omega_1\times(\Omega_2\times\Omega_3),$$

$$A_1\times A_2\times A_3=(A_1\times A_2)\times A_3=A_1\times(A_2\times A_3).$$

一般地我们约定乘积集的运算可以任意地结合.

类似于例 11, 我们可以证明

定理 10 设 $\mathscr{A}_i$ 是 Ω_i 中的 σ-代数, $i=1,2,\cdots,n,1<i_1<i_2<\cdots<i_k\cdots<n$, 我们有

$$\mathscr{A}_1\times\cdots\times\mathscr{A}_n=(\mathscr{A}_1\times\cdots\times\mathscr{A}_{i_1})\times(\mathscr{A}_{i_1+1}\times\cdots\times\mathscr{A}_{i_2})\times\cdots\times(\mathscr{A}_{i_k}\times\cdots\times\mathscr{A}_n).$$

特别

$$\mathscr{A}_1\times\cdots\times\mathscr{A}_n=(\mathscr{A}_1\times\cdots\times\mathscr{A}_k)\times(\mathscr{A}_{k+1}\times\cdots\times\mathscr{A}_n).$$

这个定理说明, 对于乘积集的任何一种结合方式所得到的乘积 σ-代数都是一样的, 更重要的是, 这个定理还提供了将任意有限维乘积空间的性质作为二维乘积空间的性质的推广而得出的可能性.

类似于例 11 可以得到

$$\mathscr{B}^{(n)}=\underbrace{\mathscr{B}^{(1)}\times\cdots\times\mathscr{B}^{(1)}}_{n\text{ 个}};\mathscr{B}_z^{(n)}=\underbrace{\mathscr{B}_z^{(1)}\times\cdots\times\mathscr{B}_z^{(1)}}_{n\text{ 个}}$$

及

$$\widetilde{\mathscr{B}}^{(n)}=\underbrace{\widetilde{\mathscr{B}}^{(1)}\times\cdots\times\widetilde{\mathscr{B}}^{(1)}}_{n\text{ 个}};\widetilde{\mathscr{B}}_z^{(n)}=\underbrace{\widetilde{\mathscr{B}}_z^{(1)}\times\cdots\times\widetilde{\mathscr{B}}_z^{(1)}}_{n\text{ 个}}.$$

习题及补充

1. 设 $\mathscr{S}$ 是 Ω 上的半集代数, 试证:

i) 若 $\{A, A_1, \cdots, A_n\} \subset \mathscr{S}$, 且 $A_k \subset A, k = 1, \cdots, n$ 两两不交, 则存在 $\{A_{n+1}, \cdots, A_s\} \subset \mathscr{S}$, 使 $A_1, \cdots, A_s$ 两两不交且 $A = \sum\limits_{k=1}^{s} A_k$.

ii) 若 $\{A_1, \cdots, A_n\} \subset \mathscr{S}$, 则在 $\mathscr{S}$ 中存在两两不交的有限集族 $B_1, \cdots, B_t$, 使每一 A_k 可以表示成某些集合 B_h 的和, 即 $A_k = \sum\limits_{h \in M_k} B_h$, 其中 $M_k \subset \{1, \cdots, t\}$.

2. 设 $\mathscr{S}$ 是 Ω 上的半集代数, $\{A, A_1, A_2, \cdots\} \subset \mathscr{S}, A \subset \bigcup\limits_{n=1}^{\infty} A_n$. 试证, 存在 $A'_n \in \mathscr{S}, n = 1, 2, \cdots$, 两两不交, 且对每一 n, 存在一 m使$A'_n \subset A_m$, 且 $A = \sum\limits_{n=1}^{\infty} A'_n$.

3. 若 $\mathfrak{E} = \left\{A_k : A_k \subset \Omega, k = 1, \cdots, n,\ \text{两两不交, 且}\ \sum\limits_{k=1}^{n} A_k = \Omega\right\}$, 试证 $\mathfrak{E}$ 上的最小集代数

$$\mathscr{F}(\mathfrak{E}) = \left\{\sum_{k \in I} A_k : I \subset \{1, \cdots, n\}\right\}.$$

4. 设 $\mathfrak{E}$ 是任意一个 Ω 的子集类, 令

$\mathfrak{E}_1 \triangleq \{A : A \in \mathfrak{E}\text{或}A^c \in \mathfrak{E}\} \cup \{\Omega, \phi\}$,

$\mathfrak{E}_2 \triangleq \left\{\bigcap\limits_{k=1}^{n} A_k : A_k \in \mathfrak{E}_1, k = 1, \cdots, n, n \geqslant 1\right\}$,

$\mathfrak{E}_3 \triangleq \left\{\sum\limits_{j=1}^{m} B_j : B_j \in \mathfrak{E}_2, j = 1, \cdots, m, \text{两两不交}, m \geqslant 1\right\}$, 则 $\mathfrak{E}_3$ 是 $\mathfrak{E}$ 上的最小集代数.

5. 设 $\mathfrak{E}$ 是直线上一切开、闭及半闭区间作成的集类; $\mathfrak{E}_1$ 是一切左开右闭的有限区间 $(a, b]$ 作成的集类; $\mathfrak{E}_2$ 是一切有限开区间 (a, b) 作成的集类. $\mathfrak{E}_3$ 是一切半无限区间 $(-\infty, a)$ 作成的集类. 试证: $\sigma(\mathfrak{E}) = \sigma(\mathfrak{E}_1) = \sigma(\mathfrak{E}_2) = \sigma(\mathfrak{E}_3) = \mathscr{B}^{(1)}$,其中$\mathscr{B}^{(1)}$ 是直线上的 Borel 域.

6. 若 $\mathfrak{E} = \{A_k : A_k \subset \Omega, k = 1, 2, \cdots, \text{两两不交}\}$, 试求$\sigma(\mathfrak{E})$.

7. 设 Ω 是一不可列集, $\mathfrak{E}$ 是 Ω 的一切有限子集、可列子集及以有限子集或可列子集为余集的子集所作成的集类, 试证 $\mathfrak{E}$ 是一个 σ-代数, 若 $\mathfrak{E}$ 是 Ω 的一切有限子集及以有限子集为余集的子集作成的集类, 则对 $\mathfrak{E}$ 可以作出什么结论?

8. 设 $\mathfrak{E}$ 是任意集类, $A \in \sigma(\mathfrak{E})$, 则有 $\mathfrak{E}$ 的一个可列子类 $\mathscr{D}$ 存在 (即 $\mathscr{D}$ 只含有 $\mathfrak{E}$ 中可列个集), 使得 $A \in \sigma(\mathscr{D})$.

9. 设 $\mathfrak{E}$ 是 Ω 的子集作成的集类, A 是 Ω 的一个非空子集, 令

$$\mathfrak{E}\cap A=\{A':A'=B\cap A,B\in\mathfrak{E}\},$$

试证: $\mathscr{F}(\mathfrak{E})\cap A$ 是以 A 为空间的集类 $\mathfrak{E}\cap A$ 上的最小集代数, $\sigma(\mathfrak{E})\cap A$ 是以 A 为空间的集类 $\mathfrak{E}\cap A$ 上的最小 σ-代数.

10. 设 A 是 Ω 的一个子集, $\mathfrak{E}$ 是 Ω 的包含 A 的一切子集所作成的集类. 试指出 $\sigma(\mathfrak{E})$ 是由哪些子集作成的.

11. 设 $\mathfrak{E}$ 是一个含有 Ω 且对对称差与可列交两种运算封闭的 Ω 的子集类, 问 $\mathfrak{E}$ 是不是一个 σ-代数?

12. 设 S 是一组集运算, 若 Ω 的非空子集类 $\mathfrak{E}$ 对 S 中的一切集运算封闭, 则称为一个 S 类. 试证:

i) 任意多个 S 类的交还是 S 类.

ii) 设 $\mathfrak{E}$ 是 Ω 的一个子集类, 则存在唯一的 $\mathfrak{E}$ 上的最小 S 类 $\mathscr{A}_0$满足 (1)$\mathscr{A}_0\supset\mathfrak{E}$; (2)若$S$类$\mathscr{A}\supset\mathfrak{E}$, 则 $\mathscr{A}\supset\mathscr{A}_0$.

13. 设 $\mathscr{A}_i=\{\Omega_i,A_i,A_i^c,\phi\},i=1,2,$ 试求$\mathscr{A}_1\times\mathscr{A}_2$.

14. 设 $(R^{(1)},\mathscr{B}^{(1)})$ 是 Borel 空间, 试证

$$\{(x,y):x=y,x,y\in R^{(1)}\}\in\mathscr{B}^{(1)}\times\mathscr{B}^{(1)}.$$

15. 设 $\mathscr{A}_i$ 是 Ω_i 上的 σ-代数, $i=1,2,3$, 试证

$$\mathscr{A}_1\times\mathscr{A}_2\times\mathscr{A}_3=(\mathscr{A}_1\times\mathscr{A}_2)\times\mathscr{A}_3.$$

16. 若 $\mathscr{A}$ 是集代数且对一切两两不交的集序列的和封闭, 则 $\mathscr{A}$ 是 σ-代数.

17. 称空间 Ω 中满足下述条件的集系为 d-系:

1) $\Omega\in\mathscr{D}$;

2) 若 $A,B\in\mathscr{D}$ 且 $A\cap B=\phi$, 则 $A+B\in\mathscr{D}$;

3) 若 $A\subset B,A,B\in\mathscr{D}$, 则 $B-A\in\mathscr{D}$.

试证: π-系上的最小 d-系等于 π-系上的最小集代数. 即若 $\mathscr{C}$ 是 π-系, 则

$$\mathscr{D}(\mathscr{C})=\mathscr{F}(\mathscr{C}).$$

其中 $\mathscr{D}(\mathscr{C})$和$\mathscr{F}(\mathscr{C})$ 分别表示 $\mathscr{C}$ 上的最小 d-系和最小集代数.

§1.4 集函数、测度与概率

在 §1.1 引言中, 我们用数 $P(A)$ 去刻划随机事件 A 发生的可能性, 称之为事件 A 的概率. 在建立了随机事件类与基本事件集合 Ω 的某一子集类的对应关系之后,

概率就是以集类为定义域的函数, 这种函数叫做集函数. 本节讨论集函数及其特殊情形测度与概率, 并把空间 Ω, Ω 中的集类 σ-代数 $\mathscr{A}$ 及定义在 $\mathscr{A}$ 上的测度 μ(或概率 P) 三者联系起来, 建立测度空间 (或概率空间) 的概念, 这是今后讨论一切问题的基础.

1.4.1　集函数与测度

在今后的讨论中不仅要用到像概率这样的集函数, 而且还要用到其他类型的集函数, 例如

1) 区间 Δ(实数集) 的长度 $l(\Delta)$;

2) 平面图形 F(实数对集) 的面积 $\delta(F)$;

3) 对于给定的函数 $f(x)$ 在区间 Δ 上的积分

$$\varphi(\Delta) \widehat{=} \int_{\Delta} f(x)dx.$$

在物理学中存在着各种类型的集函数, 例如物体 A 的质量 $m(A)$ 等. 我们先引入下列的

定义 1　设 $\mathfrak{C}$ 是 Ω 的一个子集类, 则以 $\mathfrak{C}$ 为定义域而在广义实直线 (即一切实数及 $\pm\infty$ 的全体) 上取值的函数 φ(即 $\varphi(A), A \in \mathfrak{C}$, 为实数或 $\pm\infty$) 叫做定义在 $\mathfrak{C}$ 上的一个**集函数**.

I. 若 $A \in \mathfrak{C}, B \in \mathfrak{C}, A \cup B \in \mathfrak{C}, A \cap B = \varnothing$,有

$$\varphi(A + B) = \varphi(A) + \varphi(B),$$

则称 φ 具有**可加性**, 或称 φ 为**可加集函数**;

II. 若对任意正整数 n 及任意 $A_k \in \mathfrak{C}, k = 1, \cdots, n$, 两两不交且 $\sum\limits_{k=1}^{n} A_k \in \mathfrak{C}$, 均有 $\varphi\left(\sum\limits_{k=1}^{n} A_k\right) = \sum\limits_{k=1}^{n} \varphi(A_k)$ 成立, 则称 φ 具有**有限可加性**, 或称 φ 为**有限可加集函数**;

III. 若 $A_n \in \mathfrak{C}, n = 1, 2, \cdots$ 两两不交且 $\sum\limits_{n=1}^{\infty} A_n \in \mathfrak{C}$, 则 $\varphi\left(\sum\limits_{n=1}^{\infty} A_n\right) = \sum\limits_{n=1}^{\infty} \varphi(A_n)$, 就称 φ 具有 σ-**可加性**或称 φ 为 σ-**可加集函数**;

IV. 若每一 $\varphi(A), A \in \mathfrak{C}$ 都是有限值, 则称 φ 为**有限集函数**.若对于每一 $A \in \mathfrak{C}$, 有 $\mathfrak{C}$ 中的集序列 $\{A_n\} \subset \mathfrak{C}$ 存在, 使得 $A = \bigcup\limits_{n=1}^{\infty} A_n$ 且 $\varphi(A_n), n = 1, 2, \cdots$ 有限, 则称 φ 为 σ-**有限集函数**;

V. 若集函数为有限可加的且只取非负值, 则称为**有限可加测度**.若集函数为 σ-可加的且只取非负值, 则称为**测度**(一般用 μ,ν 表示). 具有性质 $\mu(\Omega)=1$ 的测度称为**正规测度**, 亦称**概率**, 一般用 P 表示.

对于可加集函数 (包括有限可加与 σ-可加集函数), 为了要使得和数 $\Sigma\varphi(A_j)$ 永远是有意义的, 必须排除出现 $+\infty-\infty$ 形状的表达式的可能. 于是, 假使和数永远存在, 但有 A,B使得$\varphi(A)=+\infty,\varphi(B)=-\infty$, 如果定义域 $\mathfrak{E}$ 对求余运算封闭且 $\Omega\in\mathfrak{E}$, 则有 $\varphi(\Omega)=\varphi(A)+\varphi(A^c)=+\infty$,并且又有$\varphi(\Omega)=\varphi(B)+\varphi(B^c)=-\infty$, 这与函数 φ 的单值性矛盾. 因此, 对于可加集函数, 在它的值域中不能同时含有 $+\infty$和 $-\infty$, 为了避免讨论中的繁琐, 如无特别声明, 我们总是假定 φ 不取 $-\infty$ 为值. 并且还规定一个集函数至少在一个集上取有限值.

σ-可加集函数、测度与概率有很多类似的性质, 我们将在下面加以讨论.

性质 1 若 φ 是 $\mathfrak{E}$ 上可加 (或有限可加, 或 σ-可加) 集函数, $\varnothing\in\mathfrak{E}$, 则 $\varphi(\varnothing)=0$.

证 由于我们规定任何集函数至少在一个集合上取有限值, 设 $A\in\mathfrak{E},\varphi(A)$ 有限, 则当 φ 具有可加性 (或有限可加性) 时,

$$\varphi(A)=\varphi(A+\varnothing)=\varphi(A)+\varphi(\varnothing).$$

等式两端同减 $\varphi(A)$即得$\varphi(\varnothing)=0$. 当 φ 具有 σ-可加性时, 令 $A_1=A,A_n=\varnothing,n=2,3,\cdots$. 则由 σ-可加性知

$$\varphi(A)=\varphi\left(\sum_{n=1}^{\infty}A_n\right)=\varphi(A)+\varphi(\varnothing)+\varphi(\varnothing)+\cdots,\text{故必须}\varphi(\varnothing)=0. \qquad \square$$

性质 2 若 φ 是 $\mathfrak{E}$ 上 σ-可加集函数, $\varnothing\in\mathfrak{E}$, 则 φ 是有限可加集函数.

证 若 $A_k\in\mathfrak{E},k=1,\cdots,n$ 两两不交且 $\sum\limits_{k=1}^{n}A_k\in\mathfrak{E}$, 令 $A_{n+j}=\varnothing,j=1,2,\cdots$, 则 $\sum\limits_{k=1}^{\infty}A_k=\sum\limits_{k=1}^{n}A_k\in\mathfrak{E}$, 根据 σ-可加性及性质 1 有

$$\begin{aligned}\varphi\left(\sum_{k=1}^{n}A_k\right)=&\varphi\left(\sum_{k=1}^{\infty}A_k\right)=\sum_{k=1}^{\infty}\varphi(A_k)\\&=\sum_{k=1}^{n}\varphi(A_k)+\varphi(\phi)+\cdots+\varphi(\phi)+\cdots=\sum_{k=1}^{n}\varphi(A_k).\end{aligned}$$

故 φ 具有有限可加性. $\square$

根据定义, 有限可加集函数显然是可加的.

性质 3 设 φ 是 $\mathfrak{E}$ 上的可加集函数, 若 $A_k\in\mathfrak{E},\sum\limits_{i=1}^{k}A_j\in\mathfrak{E},k=1,\cdots,n$, 且 $A_1,\cdots,A_n$ 两两不交, 则

$$\varphi\left(\sum_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \varphi(A_k).$$

特别, 集代数 $\mathscr{F}$ 上的可加集函数一定是有限可加的.

应该注意的是一般集类 $\mathfrak{C}$ 上的可加集函数并不一定是有限可加的.

性质 4　设 φ 是集代数 $\mathscr{F}$ 上的可加集函数, $A \in \mathscr{F}, B \in \mathscr{F}$, 且 $A \subset B$, 则

$$\varphi(B) = \varphi(A) + \varphi(B - A).$$

若 $\varphi(A) < \infty$, 则

$$\varphi(B - A) = \varphi(B) - \varphi(A), \qquad \text{(可减性)}$$

若 μ 为有限可加测度, 则

$$\mu(A) \leqslant \mu(B). \qquad \text{(不降性)}$$

性质 5　设 μ 是半集代数 $\mathscr{S}$ 上的测度或有限可加测度, 若 $A \subset B, A, B \in \mathscr{S}$, 则

$$\mu(A) \leqslant \mu(B).$$

证　因为 $\mathscr{S}$ 是半集代数, 因而存在 $A_k \in \mathscr{S}, k = 1, \cdots, n$ 两两不交且 $A_1 = A$. 使

$$B = \sum_{k=1}^{n} A_k,$$

于是根据 μ 的有限可加性有

$$\mu(B) = \mu(A) + \sum_{k=2}^{n} \mu(A_k).$$

再由 μ 的非负性可知

$$\mu(A) \leqslant \mu(B).$$ □

性质 6　设 φ 是半集代数 $\mathscr{S}$ 上的有限可加集函数, 若 $\varphi(B) < \infty$ 且 $A \subset B$, 则 $\varphi(A) < \infty$. 特别若 $\varphi(\Omega) < \infty$, 则 φ 为有限集函数.

若有 $A_n \in \mathscr{S}, \varphi(A_n)$有限, $n = 1, 2, \cdots$ 使得 $\Omega = \bigcup_{n=1}^{\infty} A_n$, 则 φ 为 σ-有限集函数且对任何 $A \in \mathscr{S}$, 存在两两不交集序列 $\{A'_n\} \subset \mathscr{S}, \varphi(A'_n)$ 有限使得 $A = \sum_{n=1}^{\infty} A'_n$.

证　若 $A, B \in \mathscr{S}, A \subset B$ 且 $\varphi(B)$ 有限, 由性质 5 的证明知存在 $A_2, \cdots, A_n \in \mathscr{S}$, 使

$$\varphi(B)=\varphi(A)+\sum_{k=2}^{n}\varphi(A_k).$$

由于 $\varphi(B)<\infty$, 所以 $\varphi(A)<\infty$. 取 $B=\Omega$, 即知若 $\varphi(\Omega)<\infty$, 则 φ 为有限集函数.

若有 $\{A_n\}\subset\mathscr{S},\varphi(A_n)<\infty,n=1,2,\cdots$ 使 $\Omega=\bigcup_{n=1}^{\infty}A_n$, 则令 $B_n\hat{=}A_n-\left(\bigcup_{k=1}^{n-1}A_k\right)=A_n\cap\left(\bigcap_{k=1}^{n-1}A_k^c\right)$, 由于 $A_k^c,k=1,\cdots,n-1$, 为 $\mathscr{S}$ 中两两不交集的和, 故 B_n 为 $\mathscr{S}$ 中两两不交集的和, 设 $B_n=\sum_{k=1}^{j_n}A_{nk},n=1,2,\cdots$. 于是 $\{A_{nk},k=1,\cdots j_n,n=1,2,\cdots\}$ 为 $\mathscr{S}$ 中两两不交的集序列, 因为 $A_{nk}\subset B_n\subset A_n$, 故 $\varphi(A_{nk})<\infty$. 对任意 $A\in\mathscr{S},A=\sum_{n=1}^{\infty}\sum_{k=1}^{j_n}A\cap A_{nk},\varphi(A\cap A_{nk})$ 有限, $A\cap A_{nk},k=1,\cdots,j_n,n=1,2,\cdots$ 两两不交, 重新编号即得存在 $\mathscr{S}$ 中两两不交集序列 $\{A_n'\}$ 使 $A=\sum_{n=1}^{\infty}A_n'$, $\varphi(A_n')$有限, $n=1,2,\cdots$. 因而 φ 是 σ-有限集函数. □

性质 7 i) 设 μ 是集代数 $\mathscr{F}$ 上的有限可加测度, 若 $\{A_1,\cdots,A_n\}$ 是 $\mathscr{F}$ 中的有限类且 $A\in\mathscr{F},A\subset\bigcup_{k=1}^{n}A_k$, 则

$$\mu(A)\leqslant\sum_{k=1}^{n}\mu(A_k);$$

ii) 设 μ 是集代数 $\mathscr{F}$ 上的测度, 若 $A\in\mathscr{S},\{A_n\}\subset\mathscr{F},A\subset\bigcup_{n=1}^{\infty}A_n$, 则

$$\mu(A)\leqslant\sum_{k=1}^{\infty}\mu(A_k).$$

性质 7 称为半可加性和半 σ-可加性.

证 i), ii) 的证明类似, 我们只证明 ii). 令 $A_n'=A\cap A_n\in\mathscr{F}$, 则 $A=\bigcup_{n=1}^{\infty}A_n'$. 令 $A_n''=A_n'-\bigcup_{k=1}^{n-1}A_k'$, 则 $\{A_n''\}$为$\mathscr{F}$ 中两两不交集的序列且 $A=\bigcup_{n=1}^{\infty}A_n'=\sum_{n=1}^{\infty}A_n''$. 由测度的不降性及 σ-可加性及 $A_n''\subset A_n'\subset A_n$ 知

$$\mu(A)=\sum_{n=1}^{\infty}\mu(A_n'')\leqslant\sum_{n=1}^{\infty}\mu(A_n).$$ □

下面我们要研究 σ-可加集函数的连续性, 它们对测度论与概率论是很重要的, 为此我们先引进

定义 2　设 φ 是定义在集类 $\mathfrak{C}$ 上的集函数, 若 $A \in \mathfrak{C}$,对$\mathfrak{C}$ 中任何满足条件 $A_n \uparrow$ 且 $\bigcup_{n=1}^{\infty} A_n = A$ 的集序列 $\{A_n\}$, 都有

$$\lim_{n\to\infty} \varphi(A_n) = \varphi(A),$$

则称 φ 在 A**处下连续**, 若 φ 在 $\mathfrak{C}$ 中的每一 A 处下连续, 则称 φ**为下连续**.

若 $A \in \mathfrak{C}$ 且对 $\mathfrak{C}$ 中任何满足条件 $A_n \downarrow, \bigcap_{n=1}^{\infty} A_n = A$ 且有一 m 使 $\varphi(A_m) < \infty$ 的集序列 $\{A_n\}$, 都有

$$\lim_{n\to\infty} \varphi(A_n) = \varphi(A),$$

则称 φ 在 A **处上连续**, 若 φ 在 $\mathfrak{C}$ 中每一 A 处上连续, 则称 φ **为上连续**.

若 φ 在 A 处上、下连续, 则称φ **在 A 处连续**;若 φ在$\mathfrak{C}$ 中每一 A 处连续, 则称 φ **为连续**.

显然, φ 为下 (上) 连续的充分与必要条件是对 $\mathfrak{C}$ 中任何满足条件: $A_n \uparrow$ 且 $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{C}\left(\text{相应地}A_n \downarrow, \bigcap_{n=1}^{\infty} A_n \in \mathfrak{C}\text{且有一}m\text{使}\varphi(A_m) < \infty\right)$ 的集序列 $\{A_n\}$, 有

$$\lim_{n\to\infty} \varphi(A_n) = \varphi\left(\bigcup_{n=1}^{\infty} A_n\right)\left(\text{相应地, } \lim_{n\to\infty} (A_n) = \varphi\left(\bigcap_{n=1}^{\infty} A_n\right)\right).$$

下面两个定理说明 σ-可加性与连续性之间的关系.

定理 1　设 φ 是集代数 $\mathscr{F}$ 上的 σ-可加集函数, 则 φ 有限可加且连续.

证　由性质 2 知 φ 是有限可加的, 今证 φ 是连续的.

i) 设 $\{A_n\} \subset \mathscr{F}, A_n \uparrow$ 且 $\bigcup_{n=1}^{\infty} A_n \in \mathscr{F}$,若存在$m$使$\varphi(A_m) = \infty$, 则由性质 4 易知

$$\varphi\left(\bigcup_{n=1}^{\infty} A_n\right) = \infty = \lim_{n\to\infty} \varphi(A_n).$$

若对一切 $n, \varphi(A_n)$ 有限, 则

$$\bigcup_{n=1}^{\infty} A_n = A_1 + \sum_{n=2}^{\infty}(A_n - A_{n-1})$$

且 $A_1, A_n - A_{n-1}, n = 2, 3 \cdots$ 是两两不交的 $\mathscr{F}$ 中的集序列, 故由 σ-可加性及性质 4 可知

$$\begin{aligned}\varphi\left(\bigcup_{n=1}^{\infty} A_n\right) &= \varphi(A_1)+\sum_{n=2}^{\infty}\varphi(A_n-A_{n-1})\\ &= \varphi(A_1)+\lim_{N\to\infty}\sum_{n=2}^{N}\varphi(A_n-A_{n-1})\\ &= \varphi(A_1)+\lim_{N\to\infty}\sum_{n=2}^{N}[\varphi(A_n)-\varphi(A_{n-1})]\\ &= \lim_{N\to\infty}\varphi(A_N),\end{aligned}$$

故 φ 是下连续的.

ii) 设 $\{A_n\}\subset\mathscr{F}, A_n\downarrow, \bigcap_{n=1}^{\infty} A_n\in\mathscr{F}$ 且有一 m 使 $\varphi(A_m)<\infty$, 令 $A'_n=A_m-A_n, n=1,2,\cdots$, 则 $A'_n\uparrow$ 且

$$\bigcup_{n=1}^{\infty} A'_n=\bigcup_{n=1}^{\infty}(A_m-A_n)=A_m\left(\bigcap_{n=1}^{\infty} A_n\right)^c\in\mathscr{F},$$

于是由 i) 知

$$\varphi\left(\bigcup_{n=1}^{\infty} A'_n\right)=\lim_{N\to\infty}\varphi(A'_N). \tag{1}$$

但由性质 4 知当 $N\geqslant m$ 时, 由于 $\varphi(A_m)$ 有限,

$$\varphi(A'_N)=\varphi(A_m-A_N)=\varphi(A_m)-\varphi(A_N) \tag{2}$$

且

$$\begin{aligned}\varphi\left(\bigcup_{n=1}^{\infty} A'_n\right) &= \varphi\left(A_m-\bigcap_{n=1}^{\infty} A_n\right)\\ &= \varphi(A_m)-\varphi\left(\bigcap_{n=1}^{\infty} A_n\right).\end{aligned} \tag{3}$$

将 (2), (3) 代入 (1) 即得

$$\varphi(A_m)-\varphi\left(\bigcap_{n=1}^{\infty} A_n\right)=\varphi(A_m)-\lim_{N\to\infty}\varphi(A_N).$$

由于 $\varphi(A_m)$ 有限, 故

$$\varphi\left(\bigcap_{n=1}^{\infty} A_n\right)=\lim_{N\to\infty}\varphi(A_N),$$

即 φ 是上连续的. □

定理 2　设 φ 是集代数 $\mathscr{F}$ 上的有限可加集函数. 若 φ 满足下列两条件之一: (a) φ 是下连续的; (b) φ 是有限集函数且在空集 $\varnothing$ 处连续, 则 φ 是 σ-可加集函数.

证　若 $A_n \in \mathscr{F}, n=1,2,\cdots$ 两两不交且 $\sum_{n=1}^{\infty} A_n \in \mathscr{F}$,

i) 设 φ 是下连续的, 令 $B_n = \sum_{k=1}^{n} A_k$, 则 $B_n \uparrow, B_n \in \mathscr{F}, n=1,2,\cdots$, $\bigcup_{n=1}^{\infty} B_n = \sum_{n=1}^{\infty} A_n$, 于是由 φ 的下连续性及有限可加性知

$$\varphi\left(\sum_{n=1}^{\infty} A_n\right) = \varphi\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \varphi(B_n) = \lim_{n\to\infty} \varphi\left(\sum_{k=1}^{n} A_k\right)$$
$$= \lim_{n\to\infty} \sum_{k=1}^{n} \varphi(A_k) = \sum_{k=1}^{\infty} \varphi(A_k),$$

即 φ 是 σ-可加的.

ii) 设 φ 有限, 则对每一 n 来说,

$$\sum_{k=n+1}^{\infty} A_k = \sum_{k=1}^{\infty} A_k - \sum_{k=1}^{n} A_k \in \mathscr{F},$$

故

$$\varphi\left(\sum_{k=1}^{\infty} A_k\right) = \varphi\left(\sum_{k=1}^{n} A_k\right) + \varphi\left(\sum_{k=n+1}^{\infty} A_k\right)$$
$$= \sum_{k=1}^{n} \varphi(A_k) + \varphi\left(\sum_{k=n+1}^{\infty} A_k\right). \tag{4}$$

令 $B_n = \sum_{k=n+1}^{\infty} A_k$, 则 $B_n \downarrow$ 且 $\bigcap_{n=1}^{\infty} B_n = \varnothing$, 故由$\varphi$有限及在$\varnothing$ 处连续知

$$\lim_{n\to\infty} \varphi\left(\sum_{k=n+1}^{\infty} A_k\right) = \lim_{n\to\infty} \varphi(B_n) = \varphi(\varnothing) = 0.$$

在 (4) 中令 $n \to \infty$ 即得

$$\varphi\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \varphi(A_k).$$

故 φ 是 σ-可加的.　□

1.4.2 测度空间与概率场

定义 3 设 Ω 是一个集合, $\mathscr{A}$ 是 Ω 的一些子集构成的 σ-代数, μ 是 $\mathscr{A}$ 上的测度, 则称 Ω 为空间, $(\Omega,\mathscr{A})$ 为**可测空间**, $(\Omega,\mathscr{A},\mu)$ 为**测度空间**. $\mathscr{A}$ 中的元称为 $\mathscr{A}$-**可测集**或简称**可测集**.

定义 4 设 Ω 是一个基本事件集合, $\mathscr{A}$ 是 Ω 的一些子集构成的 σ-代数, P 是 $\mathscr{A}$ 上的概率, 则称 Ω 为**基本事件空间**, $(\Omega,\mathscr{A})$ 为可测空间, $(\Omega,\mathscr{A},P)$ 为**概率空间**或**概率场**.

概率是一个正规测度, 它具有测度的一切性质, 为了今后的应用, 在此重述一下概率的基本性质及由它们导出的常用性质. 设 $(\Omega,\mathscr{A},P)$ 为概率场, 则

I) 非负性. 对一切 $A\in\mathscr{A}, P(A)\geqslant 0$;

II) 正规性.$P(\Omega)=1$;

III) σ-可加性. 若 $A_n\in\mathscr{A}, n=1,2,\cdots$ 两两不交, 则$P\left(\sum\limits_{n=1}^{\infty}A_n\right)=\sum\limits_{n=1}^{\infty}P(A_n)$;

以上是概率的基本性质. 概率还具有以下性质:

IV) 有限可加性. 若 $A_k\in\mathscr{A}, k=1,\cdots,n$且两两不交, 则$P\left(\sum\limits_{k=1}^{\infty}A_k\right)=\sum\limits_{k=1}^{n}P(A_k)$;

V) 可减性. 若 $A,B\in\mathscr{A}, B\supset A$, 则

$$P(B-A)=P(B)-P(A).$$

特别, 若 $A\in\mathscr{A}$, 则

$$P(A^c)=1-P(A);$$

VI) 不降性. 若 $A,B\in\mathscr{A}, B\supset A$, 则

$$P(A)\leqslant P(B),$$

特别, 对一切 $A\in\mathscr{A}$,

$$P(A)\leqslant 1;$$

VII) 加法公式: 若 $A,B\in\mathscr{A}$, 则

$$P(A\cup B)=P(A)+P(B)-P(A\cap B). \tag{5}$$

一般来说, 若 $A_k\in\mathscr{A}, k=1,\cdots,n$, 则

$$P\left(\bigcup_{k=1}^{n}A_k\right)=\sum_{k=1}^{n}P(A_k)-\sum_{1\leqslant i<j\leqslant n}P(A_i\cap A_j)$$

$$+\cdots+(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}P(A_{i_1}\cap\cdots\cap A_{i_k})$$
$$+\cdots+(-1)^{n-1}P(A_1\cap\cdots\cap A_n). \tag{6}$$

证　i) 由于 $A\cup B=A+(B-A\cap B)$ 且 A与$(B-A\cap B)$ 不交, 由 IV) 知 $P(A\cup B)=P(A)+P(B-(A\cap B))$, 再由可减性即得 (5).

ii) 用数学归纳法证明 (6) 成立. 当 $n=2$ 时由 i) 知 (6) 成立, 设 $n-1$ 时公式成立, 则对 n 的情形由 i) 及 $n-1$ 的情形知

$$\begin{aligned}
P\left(\bigcup_{k=1}^{n}A_k\right)=&P\left(\bigcup_{k=1}^{n-1}A_k\right)+P(A_n)-P\left(\left(\bigcup_{k=1}^{n-1}A_k\right)\cap A_n\right)\\
=&P\left(\bigcup_{k=1}^{n-1}A_k\right)+P(A_n)-P\left(\bigcup_{k=1}^{n-1}(A_k\cap A_n)\right)\\
=&\left[\sum_{i=1}^{n-1}P(A_i)-\sum_{1\leqslant i_1<i_2\leqslant n-1}P(A_{i_1}\cap A_{i_2})\right.\\
&+\cdots+(-1)^{k-1}\sum_{1\leqslant i_1<\cdots<i_k\leqslant n-1}P(A_{i_1}\cap\cdots\cap A_{i_k})\\
&\left.+\cdots+(-1)^{n-2}P(A_1\cap\cdots\cap A_{n-1})\right]+P(A_n)\\
&-\left[\sum_{i=1}^{n-1}P(A_i\cap A_n)-\sum_{1\leqslant i_1<i_2\leqslant n-1}P(A_{i_1}\cap A_{i_2}\cap A_n)\right.\\
&+\cdots+(-1)^{k-2}\sum_{1\leqslant i_1<\cdots<i_{k-1}\leqslant n-1}P(A_{i_1}\cap\cdots\cap A_{i_{k-1}}\cap A_n)\\
&\left.+\cdots+(-1)^{n-2}P(A_1\cap\cdots\cap A_{n-1}\cap A_n)\right]\\
=&\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i_1<i_2\leqslant n}P(A_{i_1}\cap A_{i_2})\\
&+\cdots+(-1)^{k-1}\left[\sum_{1\leqslant i_1<\cdots<i_k\leqslant n-1}P(A_{i_1}\cap\cdots\cap A_{i_k})\right.\\
&\left.+\sum_{1\leqslant i_1<\cdots<i_{k-1}\leqslant n-1}P(A_{i_1}\cap\cdots\cap A_{i_{k-1}}\cap A_n)\right]\\
&+\cdots+(-1)^{n-1}P(A_1\cap\cdots\cap A_n).
\end{aligned}$$

此即 (6) 成立.

VIII) 连续性. 若 $\{A_n\}\subset\mathscr{A}, A_n\uparrow$, 则

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n),$$

若 $\{A_n\} \subset \mathscr{A}, A_n \downarrow$, 则

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n).$$

下面举两个例子说明如何建立概率场.

例 1 在 §1.2 例 1 中, 用 A_k 表示从某工厂的某种产品中抽出 n 件恰有 k 件合格品的事件, 令 $z_n = \{0, 1, \cdots, n\}$ 及

$$\begin{aligned} \Omega &= \{A_k : k \in z_n\}, \\ B_\Lambda &= \{A_k : k \in \Lambda\}, \Lambda \subset z_n, \end{aligned}$$

则

$$\mathscr{A} = \{B_\Lambda : \Lambda \subset z_n\}$$

是 Ω 中的 σ-代数. 再令

$$P(B_\Lambda) = \sum_{k\in\Lambda} \binom{n}{k} p^k (1-p)^{n-k}, \Lambda \subset z_n,$$

其中 p 满足条件 $0 < p < 1$. 则 P 是 $\mathscr{A}$ 上满足非负性、正规性、σ-可加性的集函数, 因而是概率, 于是对 §1.2 例 1, 我们建立了概率场 $(\Omega, \mathscr{A}, P)$.

例 2 在 §1.2 例 2 中, 用 A_k 表示某公用电话系统在时间间隔 $(a, a+t]$ 内收到用户 k 次呼叫的事件, 令

$$\begin{aligned} \Omega &= \{A_k : k \in z_+\}, \\ B_\Lambda &= \{A_k : k \in \Lambda \subset z_+\}. \end{aligned}$$

则

$$\mathscr{A} = \{B_\Lambda : \Lambda \subset z_+\}$$

是 Ω 中的一个 σ-代数. 若令

$$P(B_\Lambda) = \sum_{k\in\Lambda} \frac{(\lambda t)^k}{k!} e^{-\lambda t},$$

其中 $\lambda > 0$, 则易证 P 是 $\mathscr{A}$ 上的概率. 因而 $(\Omega, \mathscr{A}, P)$ 是一概率场.

在今后的讨论中, 我们还要建立各种概率场.

习题及补充

1. 构造一个可加但非有限可加集函数的例子.

2. 若 μ 是半集代数 $\mathscr{S}$ 上的有限可加测度 (测度), 则 μ 在 $\mathscr{S}$ 上具有半可加性 (半 σ-可加性). 且若 $\{A_1,\cdots,A_n,A\}\subset\mathscr{S}$, 每一 $A_k\subset A$ 且两两不交, 则

$$\sum_{k=1}^{n}\mu(A_k)\leqslant\mu(A).$$

(提示: 利用 §1.3 习题 1)

3. 如果 φ 是 σ-代数 $\mathscr{A}$ 上的 σ-有限 σ-可加集函数, 则 $\mathscr{A}$ 中至多只有可数多个不相交集能使 $\varphi\neq 0$.

4. 设 $(\Omega,\mathscr{A})$ 是可测空间, μ 是其上有限可加测度, 且具有半可数可加性, 即若 $A\subset\bigcup\limits_{k=1}^{\infty}A_k, A\in\mathscr{A}, A_k\in\mathscr{A}, k=1,2,\cdots$, 则有

$$\mu(A)\leqslant\sum_{k=1}^{\infty}\mu(A_k),$$

则 μ 是测度.

5. 设 μ 是集代数 $\mathscr{F}$ 上的测度, $\{A_n\}$ 是 $\mathscr{F}$ 中的集序列, 且对一切 $n\geqslant 1$, $\bigcap\limits_{k=n}^{\infty}A_k\in\mathscr{F}$, $\varliminf\limits_{n\to\infty}A_n\in\mathscr{F}$, 则

$$\mu(\varliminf_{n\to\infty}A_n)\leqslant\varliminf_{n\to\infty}\mu(A_n).$$

若对一切 $n\geqslant 1$, $\bigcup\limits_{k=n}^{\infty}A_k\in\mathscr{F}$, $\varlimsup\limits_{n\to\infty}A_n\in\mathscr{F}$ 且有一 n 存在使得 $\mu\left(\bigcup\limits_{k=n}^{\infty}A_k\right)<\infty$ 则

$$\mu(\varlimsup_{n\to\infty}A_n)\geqslant\varlimsup_{n\to\infty}\mu(A_n).$$

若 μ 是有限测度且 $\varliminf\limits_{n\to\infty}A_n=\varlimsup\limits_{n\to\infty}A_n=A$(此时称 A_n 收敛向 A, 记作 $\lim\limits_{n\to\infty}A_n=A$), 则

$$\mu(A)=\lim_{n\to\infty}\mu(A_n).$$

如果 $\Sigma\mu(A_n)<\infty$, 则

$$\mu(\varlimsup_{n\to\infty}A_n)=0.$$

6. 设 $\Omega=R$, $\mathfrak{E}$是R 中的一切子集类, 令

$$\varphi(A)=\begin{cases}+\infty, & \text{若} A\in\mathfrak{E}\ \text{且}\ A\neq\varnothing,\\ 0, & \text{若} A=\varnothing.\end{cases}$$

验证 φ 为连续集函数.

7. 设 Ω 是 [0, 1] 中的一切有理数作成的集合, $\mathfrak{C}$ 是由一切形如 $A_{a,b} = \{x: x \in \Omega, a \leqslant x \leqslant b\}, 0 \leqslant a \leqslant b \leqslant 1$, 的集合作成的类, 在 $\mathfrak{C}$ 上如下定义集函数 μ:

$$\mu(A_{a,b}) = b - a.$$

试证 μ 具有有限可加性并且上、下连续, 但 μ 不具有 σ-可加性. 这对定理 2 说明了什么?

8. 设 $(\Omega, \mathscr{A}, P)$ 是概率场, 若 $B \in \mathscr{A}, P(B) > 0$, 则对一切 $A \in \mathscr{A}$, 定义 $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$, 称为 A 在 B 之下的条件概率. 试证 $P(\cdot|B)$ 是 $\mathscr{A}$ 上的概率, 因而 $(\Omega, \mathscr{A}, P(\cdot|B))$ 是概率场.

9. 设 $A_1, A_2, \cdots$ 是概率场 $(\Omega, \mathscr{A}, P)$ 中的随机事件系, 试证

i) 若 $P(A_k) = 1, k = 1, 2, \cdots, n$, 则 $P\left(\bigcap\limits_{k=1}^{n} A_k\right) = 1$;

ii) 若 $P(A_n) = 1, n = 1, 2, \cdots$, 则 $P\left(\bigcap\limits_{n=1}^{\infty} A_n\right) = 1$;

iii) 若 $P(A_n) = 0, n = 1, 2, \cdots$, 则 $P\left(\bigcup\limits_{n=1}^{\infty} A_n\right) = 0$.

10. 设 P_0, P_1 是定义在 $(\Omega, \mathscr{A})$ 上的两个概率测度. 若对任意的 $\varepsilon > 0$, 存在 $A_\varepsilon \in \mathscr{A}$, 使得 $P_1(A_\varepsilon) > 1 - \varepsilon$, 而$P_0(A_\varepsilon) < \varepsilon$, 试证: 存在 $A \in \mathscr{A}$, 使得 $P_1(A) = 1, P_0(A) = 0$. $\left(\text{提示: 取} A = \bigcap\limits_{n=1}^{\infty} \bigcup\limits_{m=n}^{\infty} A_{\frac{1}{2^m}}.\right)$

注 满足上述性质的两个概率称为是相互奇异的.

11. 设 P', P 是定义在 $(\Omega, \mathscr{A})$ 上的两个概率测度, 若对任何使 $P(A) = 0$的$A \in \mathscr{A}$, 都有 $P'(A) = 0$, 则称 P' 对 P 连续, 记作 $P' \ll P$. 试证:

i) 设 P_1, P_2 是定义在 $(\Omega, \mathscr{A})$ 上的两个概率测度, 则有 $(\Omega, \mathscr{A})$ 上的一个概率测度 P 使 $P_1 \ll P$ 且 $P_2 \ll P$.

ii) 将 i) 推广至可数无穷多个概率测度 $P_1, P_2, \cdots$ 的情形.

12. 在上连续性的定义中, 如果 $A_n \downarrow$ 且 $\bigcap\limits_{n=1}^{\infty} A_n \in \mathfrak{C}$, 但不存在 m 使 $\varphi(A_m)$ 有限, 则不要求 $\lim\limits_{n\to\infty} \varphi(A_n) = \varphi\left(\bigcap\limits_{n=1}^{\infty} A_n\right)$ 成立. 如果不是这样定义, 则定理 1 不成立. 例如: $\Omega = \{1, 2, \cdots\}$, $\mathscr{A}$为Ω 的一切子集作成的 σ-代数, $\mu(A) = A$ 中点的个数, 易见这是测度, 取 $A_n = \{n, n+1, \cdots\}$, 显然 $A_n \downarrow \varnothing$, 但 $\mu(A_n) = \infty, n = 1, 2, \cdots$ 故

$$\lim_{n\to\infty} \mu(A_n) = +\infty \neq 0 = \mu(\varnothing).$$

这个例子说明为了使 σ-可加性与连续性一致, 必须给上连续性这样下定义.

13. 设 $\mathscr{S}$ 是 Ω 中的半集代数, μ 是 $\sigma(\mathscr{S})$ 上的 σ-有限测度, 但 μ 在 $\mathscr{S}$ 上未必是 σ-有限的, 试构造出在 $\sigma(\mathscr{S})$ 上 σ-有限, 而在 $\mathscr{S}$ 上不是 σ-有限的例子.

§1.5　测度扩张定理及测度的完全化

每当使用概率的非负性、正规性、σ-可加性时, 都要检验 $\sum_{n=1}^{\infty} A_n$ 是否在所考虑的集类之中, 这就造成了应用 σ-可加性的极大困难. 因此在理论上需要把概率定义在一个 σ-代数上. 但实际上我们往生比较容易地在一个半集代数或集代数上定义概率. 于是就产生了以下的问题: 是否可以将半集代数或集代数上的概率扩充到包含它的最小 σ-代数上? 这就是本节要解决的中心问题. 为了使得零概率集的子集仍是可测集, 还需要把概率扩充到更大的一个集类上, 这就是概率的完全化问题. 为了今后的应用, 我们讨论更广泛的问题: 测度的扩张及完全化问题.

1.5.1　半集代数上的测度扩张为其最小集代数上的测度

为了确切起见, 我们先给出

定义 1　设 $\mathfrak{C}_1 \subset \mathfrak{C}_2, \mathfrak{C}_1, \mathfrak{C}_2$ 都是 Ω 的子集类, μ_i 是定义在 $\mathfrak{C}_i$ 上的测度或有限可加测度, $i = 1, 2$. 如果对一切 $A \in \mathfrak{C}_1$ 有

$$\mu_1(A) = \mu_2(A),$$

则称 μ_2是μ_1 在 $\mathfrak{C}_2$ 上的**扩张**, 称 μ_1 是 μ_2 在 $\mathfrak{C}_1$ 上的**限制**.

我们首先把半集代数上的测度扩张到其最小集代数上, 进而阐明半集代数上测度的一些有用性质.

定理 1　设 μ 是半集代数 $\mathscr{S}$ 上的测度 (或有限可加测度), 则 μ 在 $\mathscr{F}(\mathscr{S})$ 上存在一个唯一的扩张 $\widetilde{\mu}$.

证　由 §1.3 定理 1 知, 对任何 $A \in \mathscr{F}(\mathscr{S})$, 存在正整数 n 及 $B_k \in \mathscr{S}, k = 1, \cdots, n$ 两两不交, 使 $A = \sum_{k=1}^{n} B_k$, 令

$$\widetilde{\mu}(A) = \sum_{k=1}^{n} \mu(B_k). \tag{1}$$

首先 $\widetilde{\mu}$ 的定义是合理的, 即 $\widetilde{\mu}(A)$ 与 A 的表示法无关. 事实上, 若还有 $A = \sum_{j=1}^{m} C_j, C_j \in \mathscr{S}, j = 1, \cdots m$, 两两不交, 则由 μ 在 $\mathscr{S}$ 上有限可加及 $\mathscr{S}$ 对交封闭知

$$\sum_{k=1}^{n} \mu(B_k) = \sum_{k=1}^{n} \sum_{j=1}^{m} \mu(B_k \cap C_j)$$

$$=\sum_{j=1}^{m}\sum_{k=1}^{n}\mu(B_k\cap C_j)=\sum_{j=1}^{m}\mu(A\cap C_j)=\sum_{j=1}^{m}\mu(C_j).$$

其次, 非负性显然. 对有限可加的 $\mu,\widetilde{\mu}$ 有限可加显然. 今证对测度 $\mu,\widetilde{\mu}$ 具有 σ-可加性. 设 $A=\sum_{n=1}^{\infty}B_n\in\mathscr{F}(\mathscr{S}),B_n\in\mathscr{F}(\mathscr{S}),n=1,2,\cdots$ 两两不交, 由 $\mathscr{F}(\mathscr{S})$ 的构造知

$$A=\sum_{i=1}^{l}A_i,A_i\in\mathscr{S},i=1,\cdots,l \text{ 两两不交},$$

$$B_n=\sum_{j=1}^{J_n}B_{nj},B_{nj}\in\mathscr{S},j=1,\cdots,J_n \text{ 两两不交},$$

$$n=1,2,\cdots.$$

令

$$C_{nij}=A_i\cap B_{nj},i=1,\cdots,l,j=1,\cdots,J_n,n=1,2,\cdots.$$

由 B_n 两两不交知 C_{nij} 两两不交, 且 $C_{nij}\in\mathscr{S}$. 故由 μ 的 σ-可加性及 $\widetilde{\mu}$ 的定义知

$$\begin{aligned}\widetilde{\mu}(A)&=\sum_{i=1}^{l}\mu(A_i)=\sum_{i=1}^{l}\mu\left(\sum_{n=1}^{\infty}\sum_{j=1}^{J_n}C_{nij}\right)\\&=\sum_{i=1}^{l}\sum_{n=1}^{\infty}\sum_{j=1}^{J_n}\mu(C_{nij})=\sum_{n=1}^{\infty}\sum_{j=1}^{J_n}\mu\left(\sum_{i=1}^{l}C_{nij}\right)\\&=\sum_{n=1}^{\infty}\sum_{j=1}^{J_n}\mu(B_{nj})=\sum_{n=1}^{\infty}\widetilde{\mu}(B_n).\end{aligned}$$

因而 $\widetilde{\mu}$ 在 $\mathscr{F}(\mathscr{S})$ 上的 σ-可加性获证.

最后证明唯一性. 设 $\hat{\mu}$ 为 μ 在 $\mathscr{F}(\mathscr{S})$ 上的另一扩张, 则对一切 $A=\sum_{k=1}^{n}B_k\in\mathscr{F}(\mathscr{S}),B_k\in\mathscr{S},k=1,\cdots,n$ 两两不交, 都有

$$\hat{\mu}(A)=\sum_{k=1}^{n}\hat{\mu}(B_k)=\sum_{k=1}^{n}\mu(B_k)=\widetilde{\mu}(A).$$

故定理 1 全部获证. □

半集代数上的测度还具有下列有用的性质:

性质 1 设 μ 是半集代数 $\mathscr{S}$ 上的有限可加测度, 若 $\{A,A_1,\cdots,A_n\}\subset\mathscr{S},A_k$ 两两不交且 $\sum_{k=1}^{n}A_k\subset A$, 则

$$\sum_{k=1}^{n}\mu(A_k)\leqslant\mu(A).\tag{2}$$

证　我们考虑 μ 在 $\mathscr{F}(\mathscr{S})$上的扩张$\widetilde{\mu}$, 则若令

$$A_{n+1} = A - \left(\sum_{k=1}^{n} A_k\right),$$

则 $\{A_1, \cdots, A_{n+1}\} \subset \mathscr{F}(\mathscr{S}), A_k$ 是两两不交的, 由 $\widetilde{\mu}$ 的有限可加性知

$$\widetilde{\mu}(A) = \sum_{k=1}^{n+1} \widetilde{\mu}(A_k) = \sum_{k=1}^{n} \widetilde{\mu}(A_k) + \widetilde{\mu}(A_{n+1}).$$

由于 $\widetilde{\mu}$ 的非负性可知

$$\sum_{k=1}^{n} \widetilde{\mu}(A_k) \leqslant \widetilde{\mu}(A).$$

由于 $\{A, A_1, \cdots, A_n\} \subset \mathscr{S}$ 及 $\widetilde{\mu}$ 是 μ 的扩张, 有

$$\sum_{k=1}^{n} \mu(A_k) \leqslant \mu(A).$$

□

性质 2　i) 若 μ 是半集代数 $\mathscr{S}$ 上的有限可加测度, $\{A, A_1, \cdots, A_n\} \subset \mathscr{S}$, 而且 $A \subset \bigcup\limits_{k=1}^{n} A_k$, 则

$$\mu(A) \leqslant \sum_{k=1}^{n} \mu(A_k); \qquad \text{(半可加性)} \tag{3}$$

ii) 若 μ 是半集代数 $\mathscr{S}$ 上的测度, $A, A_n \in \mathscr{S}, n = 1, 2, \cdots, A \subset \bigcup\limits_{n=1}^{\infty} A_n$, 则

$$\mu(A) \leqslant \sum_{n=1}^{\infty} \mu(A_n). \qquad \text{(半}\sigma\text{- 可加性)} \tag{4}$$

证　由定理 1 将 μ 扩张到 $\mathscr{F}(\mathscr{S})$ 上, 得出 $\widetilde{\mu}$, 则由 §1.4 性质 7 知对 μ 是有限可加测度的情形有

$$\widetilde{\mu}(A) \leqslant \sum_{k=1}^{n} \widetilde{\mu}(A_k);$$

对 μ 是测度的情形有

$$\widetilde{\mu}(A) \leqslant \sum_{n=1}^{\infty} \widetilde{\mu}(A_n).$$

但在 $\mathscr{S}$ 上 $\widetilde{\mu}$ 与 μ 一致, 故性质 2 获证.

□

1.5.2 集代数、半集代数上的测度扩张为其最小 σ-代数上的测度

由于我们已经解决了由半集代数到它所生成的最小集代数的测度扩张问题. 因而主要问题在于如何将集代数上的测度扩张到它所生成的 σ-代数上. 但为了方便, 我们宁愿直接由半集代数上的测度扩张到它的最小 σ-代数上.

定理 2 (测度扩张定理). 设 μ 是空间 Ω 中半集代数 $\mathscr{S}$ 上的测度, 则 μ 在 $\mathscr{S}$ 生成的最小 σ-代数 $\sigma(\mathscr{S})$ 上存在一个扩张. 若 μ 是 σ-有限的, 则 μ 在 $\sigma(\mathscr{S})$ 上的扩张是唯一的, 即若 μ_1, μ_2 都是 μ 在 $\sigma(\mathscr{S})$ 上的扩张, 则对任何 $B \in \sigma(\mathscr{S})$ 来说, 都有

$$\mu_1(B) = \mu_2(B).$$

这个定理的证明将由证明下面几个定理来完成, 在证明中还需要一些概念. 先引入 "外测度" 的概念.

定义 2 设 μ 是空间 Ω 中半集代数 $\mathscr{S}$ 上的测度, 对 Ω 中每个子集 A, 令

$$\mu^*(A) \widehat{=} \inf\left\{\sum_{n=1}^{\infty} \mu(A_n); A \subset \bigcup_{n=1}^{\infty} A_n, A_n \in \mathscr{S}\right\}. \tag{5}$$

称 $\mu^*(A)$ 为集 A 的**外测度**, 而集函数 μ^* 叫做**由 μ 引出的外测度**.

由于对任一 $A \subset \Omega$ 来说, 令 $A_1 = \Omega, A_n = \varnothing, n \geqslant 2$, 则 $A \subset \bigcup\limits_{n=1}^{\infty} A_n, A_n \in \mathscr{S}, n = 1, 2, \cdots$, 故由 (5) 定义的外测度对 A 是有意义的. 因此 μ^* 的定义域是 Ω 的一切子集构成的集类.

定理 3 设 μ 是空间 Ω 中半集代数 $\mathscr{S}$ 上的测度, 则由 μ 引出的外测度 μ^* 与 μ 在 $\mathscr{S}$ 上是一致的, 即

$$\mu^*(A) = \mu(A), \qquad A \in \mathscr{S}. \tag{6}$$

并且 μ^* 具有不降性及半 σ-可加性, 即

$$\mu^*(A) \leqslant \mu^*(B), \qquad A \subset B \subset \Omega; \tag{7}$$

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leqslant \sum_{n=1}^{\infty} \mu^*(A_n), \qquad A_n \subset \Omega, n = 1, 2, \cdots. \tag{8}$$

证 i) 先证明 (6) 式成立, 若 $A \in \mathscr{S}$, 因 $A \subset A$, 故由 (5) 知 $\mu^*(A) \leqslant \mu(A)$; 另一方面, 若 $A \subset \bigcup\limits_{n=1}^{\infty} A_n, A_n \in \mathscr{S}, n = 1, 2, \cdots$, 则由性质 2 知 $\mu(A) \leqslant \sum\limits_{n=1}^{\infty} \mu(A_n)$, 因此由 (5) 知

$$\mu(A) \leqslant \inf\left\{\sum_{n=1}^{\infty} \mu(A_n) : A \subset \bigcup_{n=1}^{\infty} A_n, \{A_n\} \subset \mathscr{S}\right\} = \mu^*(A),$$

故 (6) 成立.

ii) 再证 (7) 式成立. 设 $A \subset B$, 若 $\{A_n\} \subset \mathscr{S}$ 且 $B \subset \bigcup_{n=1}^{\infty} A_n$, 则 $A \subset \bigcup_{n=1}^{\infty} A_n$, 故由 (5) 知

$$\mu^*(A) \leqslant \sum_{n=1}^{\infty} \mu(A_n).$$

因此

$$\begin{aligned}\mu^*(A) &\leqslant \inf\left\{\sum_{n=1}^{\infty}\mu(A_n) : B \subset \bigcup_{n=1}^{\infty} A_n, \{A_n\} \subset \mathscr{S}\right\} \\ &= \mu^*(B).\end{aligned}$$

iii) 最后证明 μ^* 具有半 σ-可加性, 即 (8) 成立. 设 $\{A_n\}$ 是 Ω 中任一可数集系, 若 $\sum_{n=1}^{\infty} \mu^*(A_n) = +\infty$, 则 (8) 自然成立; 若 $\sum_{n=1}^{\infty} \mu^*(A_n) < \infty$, 令 ε 为任一正数, 于是对于每一 A_n 来说, 由 (5) 知, 有 $A_{nk} \in \mathscr{S}, k = 1, 2, \cdots$, 使得 $A_n \subset \bigcup_{k=1}^{\infty} A_{nk}$ 且

$$\sum_{k=1}^{\infty} \mu(A_{nk}) \leqslant \mu^*(A_n) + \frac{\varepsilon}{2^n} \tag{9}$$

成立, 因而 $\bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty}\bigcup_{k=1}^{\infty} A_{nk}$, 且由 (5) 及 (9) 知

$$\begin{aligned}\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) &\leqslant \sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\mu(A_{nk}) \leqslant \sum_{n=1}^{\infty}\left(\mu^*(A_n) + \frac{\varepsilon}{2^n}\right) \\ &= \sum_{n=1}^{\infty}\mu^*(A_n) + \varepsilon.\end{aligned}$$

因为 ε 可以任意小, 故 (8) 成立. □

外测度 μ^* 不一定在 Ω 的一切子集类上满足 σ-可加性甚至是有限可加性. 为了试图满足可加性的要求, 我们选出那些能够可加地分割任何其他集的集, 这就需要引入

定义 3 设 A 是 Ω 的子集, 且对任何 Ω 的子集 D 来说, 都有

$$\mu^*(D) = \mu^*(A \cap D) + \mu^*(A^c \cap D),$$

则称 A 为 **μ^*-可测集**.

μ^*-可测集有下述性质:

性质 3 A 是 Ω 中的 μ^*-可测集的充分与必要条件是对每一 $D \subset \Omega$ 来说,

$$\mu^*(D) \geqslant \mu^*(A \cap D) + \mu^*(A^c \cap D). \tag{10}$$

证 条件的必要性显然. 今证充分性. 因为由 (6) 知 $\mu^*(\varnothing)=0$, 故由 (8) 知 (令$A_1=A\cap D, A_2=A^c\cap D, A_n=\varnothing, n\geqslant 3$), 对每一 $D\subset\Omega$,

$$\mu^*(D)\leqslant\mu^*(A\cap D)+\mu^*(A^c\cap D).$$

因此由 (10) 知, 对每一 $D\subset\Omega$, 定义 3 中的等式成立, 故 A 是 μ^*-可测集. □

定理 4 设 μ 是空间 Ω 中半集代数 $\mathscr{S}$ 上的测度, μ^* 是由 μ 引出的外测度, 则一切 μ^*-可测集所作成的集系 $\mathscr{A}^*$ 是一个 σ-代数.

若 $\{A_n\}$ 是 $\mathscr{A}^*$ 中两两不交集序列, $A=\sum\limits_{n=1}^{\infty}A_n$, 则对任一 $D\subset\Omega$,

$$\mu^*(D\cap A)=\sum_{n=1}^{\infty}\mu^*(D\cap A_n), \tag{11}$$

μ^* 在 $\mathscr{A}^*$ 上的限制 (仍记作 μ^*) 是 $\mathscr{A}^*$ 上的测度.

证 i) 我们先来证明 $\mathscr{A}^*$ 是一集代数. 首先由 $\mu^*(\varnothing)=\mu(\varnothing)=0$, 得到对任何 $D\subset\Omega$,

$$\mu^*(D)=\mu^*(\Omega\cap D)+\mu^*(\Omega^c\cap D).$$

因而由 μ^*-可测集的定义知 $\Omega\in\mathscr{A}^*$; 其次, 若 $A\in\mathscr{A}^*$, 由于在 μ^*-可测集的定义中 A 与 A^c 是对称的, 因而 $A^c\in\mathscr{A}^*$; 第三, 若 $A\in\mathscr{A}^*, B\in\mathscr{A}^*$, 则对任何 $D\subset\Omega$, 由 μ^*-可测集的定义及 μ^* 的半 σ-可加性知

$$\begin{aligned}\mu^*(D)&=\mu^*(A\cap D)+\mu^*(A^c\cap D)\\&=\mu^*(A\cap D)+\mu^*(A^c\cap B\cap D)+\mu^*(A^c\cap B^c\cap D)\\&\geqslant\mu^*((A\cup A^cB)\cap D)+\mu^*((A\cup B)^c\cap D)\\&=\mu^*((A\cup B)\cap D)+\mu^*((A\cup B)^c\cap D).\end{aligned}$$

故由性质 3 知 $A\cup B\in\mathscr{A}^*$. 因此 $\mathscr{A}^*$ 是集代数.

ii) 再证 $\mathscr{A}^*$ 是 σ-代数. 为此只需证明：若 $\{A_n\}$ 是 $\mathscr{A}^*$ 中两两不交的集序列, 则 $\sum\limits_{n=1}^{\infty}A_n\in\mathscr{A}^*$(参看 §1.3 习题 16). 由于 $\{A_n\}$ 是 μ^*-可测集系, 故对任何 $D\subset\Omega$ 有

$$\begin{aligned}\mu^*(D)&=\mu^*(A_1\cap D)+\mu^*(A_1^c\cap D)\\&=\mu^*(A_1\cap D)+\mu^*(A_2\cap A_1^c\cap D)+\mu^*(A_2^c\cap A_1^c\cap D)\\&=\cdots\\&=\mu^*(A_1\cap D)+\mu^*(A_2\cap A_1^c\cap D)+\cdots\\&\quad+\mu^*(A_n\cap A_1^c\cap\cdots\cap A_{n-1}^c\cap D)\\&\quad+\mu^*(A_1^c\cap\cdots\cap A_n^c\cap D).\end{aligned}$$

由于 $A_n, n=1,2,\cdots$ 两两不交, 故 $A_k\cap A_{k-1}^c\cdots\cap A_1^c=A_k$, 而 $A_1^c\cap A_2^c\cap\cdots\cap A_n^c=$

$\left(\sum_{k=1}^{n} A_k\right)^c \supset \left(\sum_{k=1}^{\infty} A_k\right)^c$, 故由上式及 (7) 知

$$\mu^*(D) \geqslant \sum_{k=1}^{n} \mu^*(A_k \cap D) + \mu^*\left(\left(\sum_{k=1}^{\infty} A_k\right)^c \cap D\right).$$

在上式中令 $n \to \infty$ 并应用 (8) 即得

$$\begin{aligned}\mu^*(D) \geqslant & \sum_{k=1}^{\infty} \mu^*(A_k \cap D) + \mu^*\left(\left(\sum_{k=1}^{\infty} A_k\right)^c \cap D\right) \\ \geqslant & \ \mu^*\left(\left(\sum_{k=1}^{\infty} A_k\right) \cap D\right) + \mu^*\left(\left(\sum_{k=1}^{\infty} A_k\right)^c \cap D\right). \end{aligned} \tag{12}$$

故由性质 3 知 $\sum_{k=1}^{\infty} A_k \in \mathscr{A}^*$. 因此 $\mathscr{A}^*$ 是 σ-代数.

iii) 再次证 (11) 式成立. 设 $\{A_n\}$ 是 $\mathscr{A}^*$ 中两两不交的集序列, 则由 ii) 知, $\sum_{n=1}^{\infty} A_n \in \mathscr{A}^*$, 因而对任何 $D \subset \Omega$ 来说, $\mu^*(D) = \mu^*\left(\left(\sum_{n=1}^{\infty} A_n\right) \cap D\right) + \mu^*\left(\left(\sum_{n=1}^{\infty} A_n\right)^c \cap D\right)$, 将此式与 (12) 结合即得

$$\begin{aligned} & \mu^*\left(\left(\sum_{n=1}^{\infty} A_n\right) \cap D\right) + \mu^*\left(\left(\sum_{n=1}^{\infty} A_n\right)^c \cap D\right) \\ = & \sum_{n=1}^{\infty} \mu^*(A_n \cap D) + \mu^*\left(\left(\sum_{n=1}^{\infty} A_n\right)^c \cap D\right). \end{aligned} \tag{13}$$

用 $\left(\sum_{n=1}^{\infty} A_n\right) \cap D$ 代替 (13) 中的 D 即得 (由于 $\mu^*(\varnothing) = 0$)(11).

iv) 最后证明 μ^* 是 $\mathscr{A}^*$ 上的测度. 在 (11) 中令 $D = A$, 即得 $\left(注意 A = \sum_{n=1}^{\infty} A_n\right)$

$$\mu^*(A) = \sum_{n=1}^{\infty} \mu^*(A_n).$$

这就说明, μ^* 在 $\mathscr{A}^*$ 上的限制是 $\mathscr{A}^*$ 上的测度. □

有了以上的结果, 我们现在来完成定理 2 的证明.

定理 2 的证明. 要证 Ω 中半集代数 $\mathscr{S}$ 上的测度 μ 在 $\sigma(\mathscr{S})$ 上有扩张. 由定理 4 知只需证 $\sigma(\mathscr{S}) \subset \mathscr{A}^*$, 则 μ^* 在 $\sigma(\mathscr{S})$ 上的限制就是 μ 在 $\sigma(\mathscr{S})$ 上的一个扩张. 再由定理 4 知 $\mathscr{A}^*$ 是 σ-代数, 故由最小 σ-代数的定义知, 只需证 $\mathscr{S} \subset \mathscr{A}^*$.

若 $A \in \mathscr{S}$, 则由 μ^* 的定义知, 对任一 $D \subset \Omega$, 及任一正数 ε, 有 $\mathscr{C}$ 中的集序列 $\{A_n\}$, $\bigcup\limits_{n=1}^{\infty} A_n \supset D$, 使得

$$\mu^*(D) + \varepsilon \geqslant \sum_{n=1}^{\infty} \mu(A_n).$$

由于 $A \in \mathscr{S}$, 根据半集代数的性质, 存在 $B_k \in \mathscr{S}, k = 1, \cdots, m$, 两两不交且使 $A^c = \sum\limits_{j=1}^{m} B_j$, 于是 A_n 可以表示成 $\mathscr{S}$ 中 $m+1$ 个集的不交并, 即

$$A_n = A \cap A_n + \sum_{j=1}^{m} (B_j \cap A_n),$$

因而 $\mu(A_n) = \mu(A \cap A_n) + \sum\limits_{j=1}^{m} \mu(B_j \cap A_n)$. 而 $\bigcup\limits_{n=1}^{\infty} (A \cap A_n) \supset D \cap A$, $\bigcup\limits_{n=1}^{\infty} \left(\sum\limits_{j=1}^{m} B_j \cap A_n \right) \supset D \cap A^c$. 于是

$$\begin{aligned} \mu^*(D) + \varepsilon &\geqslant \sum_{n=1}^{\infty} \mu(A \cap A_n) + \sum_{n=1}^{\infty} \sum_{j=1}^{m} \mu(B_j \cap A_n) \\ &\geqslant \mu^*(A \cap D) + \mu^*(A^c \cap D). \end{aligned}$$

由于 ε 可以任意小, 因而 A 是 μ^*-可测集, 即 $A \in \mathscr{A}^*$, 因而 $\mathscr{S} \subset \mathscr{A}^*$. 这样, 定理 2 的第一部分获证.

现在证明定理 2 的后一部分. 由于 μ 在 $\mathscr{S}$ 上是 σ-有限的, 由 §1.4 性质 6 知, 存在 $D_n \in \mathscr{S}, n = 1, 2, \cdots$ 两两不交, $\sum\limits_{n=1}^{\infty} D_n = \Omega$ 且 $\mu(D_n) < \infty, n = 1, 2, \cdots$. 若 μ_1, μ_2 都是 μ 在 $\sigma(\mathscr{S})$ 上的扩张, 我们先证明对于任意 $A \in \sigma(\mathscr{S})$ 及任意正整数 n 有

$$\mu_1(A \cap D_n) = \mu_2(A \cap D_n). \tag{14}$$

为此, 对任意给定的 n, 令

$$\mathfrak{M} = \{A : A \in \sigma(\mathscr{S}), \mu_1(A \cap D_n) = \mu_2(A \cap D_n)\}.$$

显然 $\mathscr{S} \subset \mathfrak{M}$. 再由定理 1 知 μ_1, μ_2 在 $\mathscr{F}(\mathscr{S})$ 上的限制应是一致的, 因此, $\mathscr{F}(\mathscr{S}) \subset \mathfrak{M}$, 所以只要证明了 $\mathfrak{M}$ 是单调类, 即知 $\mathfrak{M} = \sigma(\mathscr{F}(\mathscr{S})) = \sigma(\mathscr{S})$, 从而 (14) 获证.

现在验证 $\mathfrak{M}$ 是一单调类: 设 $A_k \in \mathfrak{M}, k = 1, 2, \cdots A_k \uparrow$, 则 $\mu_1(A_k \cap D_n) = \mu_2(A_k \cap D_n), n = 1, 2, \cdots$. 由于测度的下连续性, 我们有

$$\mu_1\left(\left(\bigcup_{k=1}^{\infty} A_k\right)\cap D_n\right)=\lim_{k\to\infty}\mu_1(A_k\cap D_n)=\lim_{k\to\infty}\mu_2(A_k\cap D_n)$$
$$=\mu_2\left(\left(\bigcup_{k=1}^{\infty} A_k\right)\cap D_n\right).$$

故 $\bigcup_{k=1}^{\infty} A_k\in\mathfrak{M}$; 又设 $A_k\in\mathfrak{M}, k=1,2,\cdots,A_k\downarrow$, 注意到 $\mu(D_n)$ 有限, 因而 $\mu_i(A_k\cap D_n)\leqslant\mu_i(D_n)=\mu(D_n)<\infty, i=1,2$, 利用测度的上连续性,

$$\mu_1\left(\left(\bigcap_{k=1}^{\infty} A_k\right)\cap D_n\right)=\lim_{k\to\infty}\mu_1(A_k\cap D_n)=\lim_{k\to\infty}\mu_2(A_k\cap D_n)$$
$$=\mu_2\left(\left(\bigcap_{k=1}^{\infty} A_k\right)\cap D_n\right).$$

即 $\bigcap_{k=1}^{\infty} A_k\in\mathfrak{M}$. 因此 $\mathfrak{M}$ 是单调类.(14) 获证.

利用 (14) 及测度的 σ-可加性, 对任意的 $A\in\sigma(\mathscr{S})$,

$$\mu_1(A)=\sum_{n=1}^{\infty}\mu_1(A\cap D_n)=\sum_{n=1}^{\infty}\mu_2(A\cap D_n)=\mu_2(A).$$

因此 μ_1 与 μ_2 在 $\sigma(\mathscr{S})$ 上一致. □

至此测度扩张定理全部证完.

对概率有以下直接推论.

推论　若 $\mathscr{S}$ 是 Ω 中的半集代数, P 是 $\mathscr{S}$ 上的概率, 则有唯一的概率场 $(\Omega,\sigma(\mathscr{S}),P^*)$ 存在, 使得对任何 $A\in\mathscr{S}$ 有

$$P^*(A)=P(A).$$

可以构造出当 μ 不是 σ-有限时扩张不唯一的例子 (参看习题 2).

测度扩张定理表明, 只要在一个半集代数上确定了 σ-有限测度, 那就在 $\sigma(\mathscr{S})$ 上自然地确定了一个测度, 这给建立测度带来了很大的方便.

例 1　设 $A_n, n=1,2,\cdots$, 两两不交且 $\sum_{n=1}^{\infty}A_n=\Omega$(称 $\{A_n\}$ 为 Ω 的一个可数分划), 则 $\mathscr{S}=\left\{A_n,\sum_{k=n}^{\infty}A_k:n\geqslant 1\right\}\cup\{\Omega,\varnothing\}$ 是 Ω 中的半集代数. 在 $\mathscr{S}$ 上定义

$$\mu(A_n)=q_n\geqslant 0,\quad n=1,2,\cdots,$$
$$\mu\left(\sum_{k=n}^{\infty}A_k\right)=\sum_{k=n}^{\infty}q_k,\quad n=1,2,\cdots,$$
$$\mu(\Omega)=\sum_{k=1}^{\infty}q_k,\quad \mu(\varnothing)=0.$$

则 μ 是 $\mathscr{S}$ 上的测度, 可以自然地扩张到 $\sigma(\mathscr{S})$ 上, 且若 $q_n, n \geqslant 1$ 皆有限, 扩张还是唯一的.

我们还将看到, 由 $\mathscr{S}^{(1)}$ 上的测度自然地扩充到一切 Borel 集类 $\mathscr{B}^{(1)}$ 上的例子.

1.5.3 测度的完全化

上一小节, 我们借助于外测度证明了在半集代数 $\mathscr{S}$ 上的测度能够扩张到用它生成的最小 σ-代数上. 在证明的过程中, 我们看到还可以把它扩张到 μ^*-可测集类 $\mathscr{A}^*$ 上, $\mathscr{A}^* \supset \sigma(\mathscr{S})$, 且 $\mathscr{A}^*$ 是一 σ-代数. 自然要问: $\mathscr{A}^*$ 同 $\sigma(\mathscr{S})$ 相比, 究竟大多少? 二者之间有什么关系? 这个问题的解决涉及到测度空间的完全化, 我们先引入

定义 4 设 $(\Omega, \mathscr{A}, \mu)$ 是一个测度空间, 称 Ω 的一个子集 B 为 **μ-零集**, 如果存在 $A \in \mathscr{A}$, 使得 $B \subset A$ 且 $\mu(A) = 0$. 称此空间为**完全的**, 如果每一 μ-零集 $B \in \mathscr{A}$.

易证, μ-零集的类对集合的可数并是封闭的.

定理 5 设 $(\Omega, \mathscr{A}, \mu)$ 是一测度空间, 令 $\bar{\mathscr{A}} = \{A \triangle N : A \in \mathscr{A}, N\text{是}\mu\text{-零集}\}$, 则 $\bar{\mathscr{A}}$ 是一 σ-代数. 若在 $\bar{\mathscr{A}}$ 上定义

$$\overline{\mu}(A \triangle N) = \mu(A),$$

其中 $A \in \mathscr{A}, N$ 是 μ-零集, 则 $\overline{\mu}$ 的定义是合理的且 $(\Omega, \bar{\mathscr{A}}, \overline{\mu})$ 是一完全测度空间 (称为 $(\Omega, \mathscr{A}, \mu)$ 的**完全化**).

证 如果 $A \in \mathscr{A}, N \subset B, B \in \mathscr{A}, \mu(B)=0$, 则利用等式

$$A \cup N = (A - B) \triangle (B \cap (A \cup N)), \tag{15}$$

$$A \triangle N = (A - B) \cup (B \cap (A \triangle N)), \tag{16}$$

可以证明

$$\bar{\mathscr{A}} = \{A \cup N : A \in \mathscr{A}, N\text{是}\mu\text{-零集}\}. \tag{17}$$

因此, $\bar{\mathscr{A}}$ 对可数并封闭, 又因为 $(A \triangle N)^c = A^c \triangle N$, 于是 $\bar{\mathscr{A}}$ 对求余运算封闭, 这样 $\bar{\mathscr{A}}$ 就是一 σ-代数.

为了证明 $\overline{\mu}$ 的定义是合理的, 设 $A_1 \triangle N_1 = A_2 \triangle N_2$, 其中 $A_1, A_2 \in \mathscr{A}, N_1, N_2$ 是 μ-零集. 容易证明对称差运算满足交换律和结合律, 因而

$$(A_1 \triangle A_2) \triangle (N_1 \triangle N_2) = (A_1 \triangle N_1) \triangle (A_2 \triangle N_2) = \varnothing,$$

即

$$A_1 \triangle A_2 = N_1 \triangle N_2.$$

由于 $N_1 \triangle N_2 \subset N_1 \cup N_2$, 故 $N_1 \triangle N_2$ 是 μ-零集, 于是 $\mu(A_1 \triangle A_2) = 0$. 因为 $A_1 = A_1 \cap A_2 + A_1 \cap A_2^c$, 故

$$\mu(A_1) = \mu(A_1 \cap A_2),$$

同理

$$\mu(A_2) = \mu(A_1 \cap A_2).$$

于是 $\mu(A_1) = \mu(A_2)$, 即

$$\overline{\mu}(A_1 \triangle N_1) = \overline{\mu}(A_2 \triangle N_2).$$

$\overline{\mu}$ 定义的合理性获证. 再由 (15) 知

$$\overline{\mu}(A \cup N) = \mu(A).$$

利用 (17) 及 μ 的 σ-可加性易证 $\overline{\mu}$ 在 $\bar{\mathscr{A}}$ 上是 σ-可加的.

最后证明 $(\Omega, \bar{\mathscr{A}}, \overline{\mu})$ 是完全测度空间. 若 $B \subset \Omega$, 且存在一 $A \in \bar{\mathscr{A}}$ 使 $B \subset A$ 且 $\overline{\mu}(A) = 0$, 则存在 $A_1 \in \mathscr{A}, \mu(A_1) = 0$及$N_1$是$\mu-$零集使$A = A_1 \triangle N_1$, 因而存在 $B_1 \in \mathscr{A}, N_1 \subset B_1, \mu(B_1) = 0$.因此,$B \subset A_1 \cup B_1 \in \mathscr{A}, \mu(A_1 \cup B_1) = 0$, 故 B 是 μ-零集, $B = \varnothing \triangle B \in \bar{\mathscr{A}}$. 这就完成了定理的证明. □

定理 6　若 μ 是半集代数 $\mathscr{S}$ 上的测度, μ^* 是在 Ω 的子集类上由 μ 引出的外测度, 若 $A \subset \Omega$, 并且 $\mu^*(A) < \infty$, 则存在一个集 $B \in \sigma(\mathscr{S})$, 使得 (i)$A \subset B$,(ii)$\mu^*(A) = \mu^*(B)$,(iii) 对一切 $C \subset B - A$ 且 $C \in \sigma(\mathscr{S})$, 都有 $\mu^*(C) = 0$.(称集 B 为集 A 的可测覆盖.)

证　对每一正整数 n, 存在 $\mathscr{S}$ 中的集序列 $\{F_{nk}\}_{k \geqslant 1}$, 使得

$$A \subset \bigcup_k F_{nk}, \qquad \mu^*(A) \leqslant \sum_k \mu(F_{nk}) \leqslant \mu^*(A) + \frac{1}{n}.$$

若令 $B_n = \bigcup\limits_k F_{nk}, n = 1, 2, \cdots$, 因为在 $\mathscr{S}$ 上 $\mu^* = \mu$, 且 μ^* 在 $\sigma(\mathscr{S})$ 上是测度, 于是

$$\mu^*(A) \leqslant \mu^*(B_n) \leqslant \sum_k \mu(F_{nk}) \leqslant \mu^*(A) + \frac{1}{n}.$$

令 $B = \bigcap\limits_{n=1}^{\infty} B_n$, 则 $B \in \sigma(\mathscr{S})$, 且 $B \supset A$, 因此对一切 $n \geqslant 1$,

$$\mu^*(A) \leqslant \mu^*(B) \leqslant \mu^*(B_n) \leqslant \mu^*(A) + \frac{1}{n},$$

令 $n \to \infty$, 即得 $\mu^*(A) = \mu^*(B)$. 如果 $C \subset B - A$ 且 $C \in \sigma(\mathscr{S})$, 则 $A \subset B - C$ 且 $\mu^*(A) \leqslant \mu^*(B - C) = \mu^*(B) - \mu^*(C) = \mu^*(A) - \mu^*(C)$. 因此 $\mu^*(C) = 0$.证毕.□

定理 7 设 μ 是半集代数 $\mathscr{S}$ 上的 σ-有限测度, μ^* 是由 μ 引出的外测度, $\mathscr{A}^*$ 是一切 μ^* 可测集作成的 σ-代数, 则测度空间 $(\Omega,\mathscr{A}^*,\mu^*)$ 是 $(\Omega,\sigma(\mathscr{S}),\mu)$ 的完全化.(此处将 μ 在 $\sigma(\mathscr{S})$ 上的扩张仍记作 μ.)

证 设 $\bar{\mathscr{A}}$ 是由完全化 $(\Omega,\sigma(\mathscr{S}),\mu)$ 得到的 σ-代数. 只要证明了 $\bar{\mathscr{A}}=\mathscr{A}^*$, 则由 (17) 知对任何 $B\in\mathscr{A}^*$, 必有 $A\in\sigma(\mathscr{S}),N$ 是 μ-零集,使得$B=A\cup N$, 因为 N 是外测度为零的 μ^*-可测集, μ^* 是 $\mathscr{A}^*$ 上的测度, 故 $\mu^*(B)=\mu^*(A)=\mu(A)$, 这就证明了 $(\Omega,\mathscr{A}^*,\mu^*)$ 是 $(\Omega,\sigma(\mathscr{S}),\mu)$ 的完全化.

往证 $\bar{\mathscr{A}}=\mathscr{A}^*$. 由于

$$\bar{\mathscr{A}}=\{A\cup N:A\in\sigma(\mathscr{S}),N\text{是}\mu\text{-零集}\},$$

对任何 $A\cup N\in\bar{\mathscr{A}},N$ 是 μ-零集, 故必为 μ^*-可测集, $A\in\sigma(\mathscr{S})\subset\mathscr{A}^*$, 由于$\mathscr{A}^*$是$\sigma$-代数, 故 $A\cup N\in\mathscr{A}^*$因此$\bar{\mathscr{A}}\subset\mathscr{A}^*$. 再证 $\mathscr{A}^*\subset\bar{\mathscr{A}}$. 设 $A\in\mathscr{A}^*,\mu^*(A)<\infty$, 令 B 是 A 的可测覆盖且 C 是 $B-A$ 的可测覆盖. 因为 μ^* 是 $\mathscr{A}^*$ 上的测度, 于是有

$$\mu^*(B-A)=\mu^*(B)-\mu^*(A)=0.$$

因此 $\mu^*(C)=0$. 显然 $A=(B-C)+(A\cap C)$. 因为 $B-C\in\sigma(\mathscr{S}),A\cap C\subset C,C\in\sigma(\mathscr{S})$ 且 $\mu^*(C)=0$, 故 $A\cap C$ 是 μ-零集, 因而 $A\in\bar{\mathscr{A}}$. 最后由 μ 是 σ-有限的, 因而 μ^* 是 σ-有限的, 从而对一切 $A\in\mathscr{A}^*$ 必有 $A_n\in\mathscr{A}^*,n=1,2,\cdots$ 两两不交 $\mu^*(A_n)<\infty$,使$A=\sum\limits_{n=1}^{\infty}A_n$, 由于前证结果知每一 $A_n\in\bar{\mathscr{A}}$, 再由 $\bar{\mathscr{A}}$ 是 σ-代数知 $A=\sum\limits_{n=1}^{\infty}A_n\in\bar{\mathscr{A}}$. 从而有 $\bar{\mathscr{A}}=\mathscr{A}^*$. □

还容易证明

定理 8 设 μ 是半集代数 $\mathscr{S}$ 上的 σ-有限测度, 则测度 μ 在 μ^*-可测集类 $\mathscr{A}^*$ 上的扩张是唯一的. 即若在 $\mathscr{S}$ 上 $\mu=\mu'$, 则在 $\mathscr{A}^*$ 上 $\mu=\mu'$.

下面的定理说明 μ^*-可测集与 $\mathscr{S}$ 的最小集代数 $\mathscr{F}(\mathscr{S})$ 中的集的关系.

定理 9 设 μ 是半集代数 $\mathscr{S}$ 上的测度, 则对一切使 $\mu^*(A)<\infty$的μ^*- 可测集 A 和任一 $\varepsilon>0$, 存在一集 $A_\varepsilon\in\mathscr{F}(\mathscr{S})$, 使 $\mu^*(A\triangle A_\varepsilon)<\varepsilon$.

证 任给 $\varepsilon>0$, 存在 $\mathscr{S}$ 中集序列 $\{B_n\}$, 使 $\bigcup\limits_{n=1}^{\infty}B_n\supset A$ 且

$$\mu^*(A)\leqslant\sum_{n=1}^{\infty}\mu^*(B_n)<\mu^*(A)+\frac{\varepsilon}{2},\tag{18}$$

因为 $\mu^*(A)<\infty$,所以$\sum\limits_{n=1}^{\infty}\mu^*(B_n)$ 收敛. 故存在 n_0 使

$$\sum_{n>n_0} \mu^*(B_n) < \frac{\varepsilon}{2}. \tag{19}$$

令 $A_\varepsilon = \bigcup_{n=1}^{n_0} B_n, B_\varepsilon = \bigcup_{n>n_0} B_n$, 由于 μ^* 是 σ-可加的, 则由 (18), (19) 知

$$\mu^*(B_\varepsilon) < \frac{\varepsilon}{2}, \tag{20}$$

$$\mu^*((A_\varepsilon \cup B_\varepsilon) - A) < \frac{\varepsilon}{2}. \tag{21}$$

由于 $A - A_\varepsilon \subset B_\varepsilon$ 且 $A_\varepsilon - A \subset (A_\varepsilon \cup B_\varepsilon) - A$, 利用 (20), (21) 可知 $\mu^*(A \triangle A_\varepsilon) < \varepsilon$. 由于对一切 $n, B_n \in \mathscr{S}$, 因而 $A_\varepsilon \in \mathscr{F}(\mathscr{S})$. □

1.5.4 直线上的 Lebesgue 测度

本节最后考虑一个关于测度扩张和完全化的十分重要而典型的例子, 讨论它的方法在今后会要用到.

$\mathscr{S}^{(1)}$ 表示 §1.3 例 3 中一切左闭右开区间类 (包知形如 $(-\infty, a), [b, \infty)$ 的无穷空间), 它是一个半集代数, 若在 $\mathscr{S}^{(1)}$ 上定义集函数

$$\begin{aligned} &\mu([a,b)) = b - a, \quad a, b \in R, \\ &\mu((-\infty, b)) = \mu([a, \infty)) = \mu((-\infty, \infty)) = +\infty, \end{aligned} \tag{22}$$

则 μ 是 $\mathscr{S}^{(1)}$ 上的非负集函数. 这样定义的集函数具有下述诸性质:

性质 4 设 $A_1, A_2, \cdots, A_n$ 是 $\mathscr{S}^{(1)}$ 中两两不相交的集, 且 $A_0 \in \mathscr{S}^{(1)}, A_k \subset A_0, k = 1, \cdots, n$, 则

$$\sum_{k=1}^{n} \mu(A_k) \leqslant \mu(A_0). \tag{23}$$

证 令 $A_k = [a_k, b_k), k = 0, 1, \cdots, n$. 当 $a_0 = -\infty$或$b_0 = +\infty$ 时, 性质显然成立, 故不失一般性, 可以假设 $a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n$ 且 $-\infty < a_0 \leqslant b_0 < +\infty$, 于是由假设知

$$a_0 \leqslant a_1 \leqslant b_1 \leqslant a_2 \leqslant b_2 \leqslant \cdots \leqslant a_n \leqslant b_n \leqslant b_0,$$

因此

$$\begin{aligned} \sum_{k=1}^{n} \mu(A_k) &= \sum_{k=1}^{n} (b_k - a_k) \leqslant \sum_{k=1}^{n} (b_k - a_k) \\ &\quad + \sum_{k=1}^{n-1} (a_{k+1} - b_k) \\ &= b_n - a_1 \leqslant b_0 - a_0 = \mu(A_0). \end{aligned}$$

□

性质 5 设闭区间 $B_0 = [a_0, b_0]$(a_0, b_0 为实数) 被 n 个开区间 $G_k = (a_k, b_k), k = 1, \cdots, n$ 的并所包含, 则

$$b_0 - a_0 < \sum_{k=1}^{n}(b_k - a_k).$$

证 由假设知有一 $k_1(1 \leqslant k_1 \leqslant n)$ 使得 $a_0 \in G_{k_1}$, 若 $b_{k_1} \leqslant b_0$, 则 $b_{k_1} \in [a_0, b_0]$, 故由假设知有一 $k_2, 1 \leqslant k_2 \leqslant n$, 使得 $b_{k_1} \in G_{k_2}$. 一般来说, 若已选 $k_1, \cdots, k_l$ 使得 $b_{k_1} \in G_{k_2}, \cdots, b_{k_{l-1}} \in G_{k_l}, b_{k_1} \leqslant b_0$, 则由假设知可选一 $k_{l+1}(1 \leqslant k_{l+1} \leqslant n)$ 使 $b_{k_l} \in G_{k_{l+1}}$, 由于 $\bigcup\limits_{k=1}^{n} G_k \supset B_0$, 故可在 $1, 2\cdots, n$ 中选取 $k_1, \cdots, k_m$ 使得

$$a_{k_1} < a_0 < b_{k_1}, a_{k_{l+1}} < b_{k_l} < b_{k_{l+1}}, l = 1, \cdots, m-1,$$

$$a_{k_m} < b_0 < b_{k_m}.$$

因此有

$$\begin{aligned} b_0 - a_0 < b_{km} - a_{k_1} &= b_{k_1} - a_{k_1} + \sum_{l=1}^{m-1}(b_{k_{l+1}} - b_{k_l}) \\ &< \sum_{l=1}^{m}(b_{k_l} - a_{k_l}) \leqslant \sum_{k=1}^{n}(b_k - a_k). \end{aligned}$$ □

性质 6 设 $A_n \in \mathscr{S}^{(1)}, n = 0, 1, \cdots$, 且

$$A_0 \subset \bigcup_{n=1}^{\infty} A_n,$$

则

$$\mu(A_0) \leqslant \sum_{n=1}^{\infty} \mu(A_n). \tag{24}$$

证 令 $A_n = [a_n, b_n), n = 0, 1, \cdots$, 若$a_0 = b_0$, 则 (24) 显然成立, 今考虑 $a_0 < b_0$ 的情形.

i) 若 $-\infty < a_0 < b_0 < +\infty$, 设 ε 是满足条件 $0 < \varepsilon < b_0 - a_0$ 的任意数, 令

$$\begin{aligned} B_0 &= [a_0, b_0 - \varepsilon], \\ G_n &= \left(a_n - \frac{\varepsilon}{2^n}, b_n\right), \qquad n = 1, 2, \cdots, \end{aligned}$$

则

$$B_0 \subset A_0 \subset \bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} G_n.$$

由于 B_0 是闭区间, G_n 是开区间, 故由有限覆盖定理知, 可从 $\{G_n\}$ 中选取有限个开区间 $G_1, \cdots, G_{n_0}$, 使得

$$B_0 \subset \bigcup_{k=1}^{n_0} G_k,$$

于是由性质 5 知

$$\mu(A_0)-\varepsilon = b_0-a_0-\varepsilon \leqslant \sum_{k=1}^{n_0}\left(b_k-a_k+\frac{\varepsilon}{2^k}\right)$$
$$\leqslant \sum_{n=1}^{\infty}(b_n-a_n)+\sum_{n=1}^{\infty}\frac{\varepsilon}{2^n}=\sum_{n=1}^{\infty}\mu(A_n)+\varepsilon.$$

由于 ε 的任意性, 性质 6 对这种情形成立.

ii) $a_0=-\infty, b_0<+\infty$, 设$B_N=[-N,b_0)$, 由 i) 知

$$b_0+N=\mu(B_N)\leqslant\sum_{n=1}^{\infty}\mu(A_n)$$

对一切自然数 N 成立. 令 $N\to\infty$, 即得 $\sum\limits_{n=1}^{\infty}\mu(A_n)=+\infty$, 故 $\mu(A_0)\leqslant\sum\limits_{n=1}^{\infty}\mu(A_n)$.

至于 $b_0=+\infty, a_0$ 为实数或 $b_0=+\infty, a_0=-\infty$ 的情形用同样的讨论方法知 (24) 成立. □

定理 10

i) 由 (22) 定义的 μ 是 $\mathscr{S}^{(1)}$ 上的 σ-有限测度.

ii) 在 $\mathscr{B}^{(1)}(=\sigma(\mathscr{S}^{(1)}))$ 上存在唯一的测度 $\overline{\mu}$ 使当 $A\in\mathscr{S}^{(1)}$ 时 $\overline{\mu}(A)=\mu(A)$.

iii) 由 μ 引出的外测度 μ^* 是一切 μ^*-可测集上的测度. μ^*-可测集称为 Lebesgue 集, 以 $\mathscr{B}_{\Lambda}^{(1)}$表示一切 Lebesgue 集的类. 其上测度 μ^* 称为 Lebesgue测度, 简称 L-测度, 用 Λ 表示, 则 $(R^{(1)},\mathscr{B}_{\Lambda}^{(1)},\Lambda)$ 是 $(R^{(1)},\mathscr{B}^{(1)},\overline{\mu})$ 的完全化.

证　只需证明 i), 至于 ii), iii) 则由 i) 及 $\mathscr{S}^{(1)}$ 是半集代数应用测度扩张定理及定理 7 直接得到.

设 $\{A_n\}$ 是 $\mathscr{S}^{(1)}$ 中两两不交的集序列, 且 $A=\sum\limits_{n=1}^{\infty}A_n\in\mathscr{S}^{(1)}$, 则由性质 4 知, 对任何 $n\geqslant 1$ 来说,

$$\sum_{k=1}^{n}\mu(A_k)\leqslant\mu(A).$$

令 $n\to\infty$ 即得

$$\sum_{k=1}^{\infty}\mu(A_k)\leqslant\mu(A).$$

故由性质 6 知

$$\mu(A)=\sum_{n=1}^{\infty}\mu(A_n).$$

即 μ 是 σ-可加的, 至于 μ 的非负性及 σ-有限性则是显然的. □

习题及补充

1. 设 μ 和 ν 是可测空间 $(\Omega,\mathscr{A})$ 上的两个有限测度, 又设 $\mathfrak{C}\subset\mathscr{A}$ 是包含 Ω 并生成 $\mathscr{A}$(即 $\sigma(\mathfrak{C})=\mathscr{A}$) 的 π-系, 如果 μ 和 ν 在 $\mathfrak{C}$ 上重合, 则它们在 $\mathscr{A}$ 上重合. 若 μ 和 ν 是 σ-有限测度结果如何? (提示: 证明 $\{A:\mu(A)=\nu(A),A\in\mathscr{A}\}$ 是包含 $\mathfrak{C}$ 的 λ-系).

2. 在测度扩张定理中半集代数一般不能换成结构更简单的集类; 条件 σ-有限除去, 定理的唯一性未必成立. 这可由下面例子看到:

i) 设 $\Omega=\{a,b,c,d\}$, $\mathfrak{C}=\{\{a,b\},\ \{c,d\},\ \{a,c\},\ \{b,d\},\Omega,\varnothing\}$ 令

$$\mu_1(\{a\})=\mu_1(\{d\})=\mu_2(\{b\})=\mu_2(\{c\})=1;$$
$$\mu_1(\{b\})=\mu_1(\{c\})=\mu_2(\{a\})=\mu_2(\{d\})=2;$$
$$\mu_i(\{x,y\})=\mu_i(\{x\})+\mu_i(\{y\}),i=1,2,\{x,y\}\in\mathfrak{C},$$
$$\mu_i(\varnothing)=0,\mu_i(\Omega)=6,i=1,2.$$

试证: $\mathfrak{C}$ 不是半集代数, 在 $\mathfrak{C}$ 上 $\mu_1=\mu_2$ 且都是 σ-可加的, 但在 $\sigma(\mathfrak{C})$ 上 $\mu_1\neq\mu_2$.

ii) 设 $\Omega=(-\infty,\infty),\mathscr{C}=\{[a,b),-\infty\leqslant a\leqslant b\leqslant+\infty$, 当 $a=-\infty$ 时 $[a,b)$ 理解为 $(-\infty,b)\}$, 令

$$\mu_1(A)=A\text{中点的个数},\quad A\in\sigma(\mathscr{C});$$
$$\mu_2(A)=2\mu_1(A),\quad A\in\sigma(\mathscr{C}).$$

试证: μ_1 和 μ_2 是 $\sigma(\mathscr{C})$ 上的测度, 但不是 σ-有限的, 在 $\mathscr{C}$ 上 $\mu_1=\mu_2$, 但在 $\sigma(\mathscr{C})$ 上 $\mu_1\neq\mu_2$.

3. 设 $\mathscr{F}$ 是 Ω 上的集代数, μ_1,μ_2 是 $\sigma(\mathscr{F})$ 上的测度且在 $\mathscr{F}$ 上 σ-有限, 设 $A\in\sigma(\mathscr{F})$ 且 $\mu_1(A)<\infty,\mu_2(A)<\infty$, 试证: 对任意的 $\varepsilon>0$, 存在共同的集 $A_\varepsilon\in\mathscr{F}$, 使得

$$\mu_i(A\,\triangle\,A_\varepsilon)<\varepsilon,\qquad i=1,2$$

同时成立. 若去掉 μ 在 $\mathscr{F}$ 上 σ-有限的条件, 此命题还成立吗? (提示: 考虑 $\mu_1+\mu_2$ 仍是 $\sigma(\mathscr{F})$ 上的测度.)

4. 证明一切外测度为零的集都是 μ^*-可测集.

5. 设 μ 为集代数 $\mathscr{F}$ 上的测度, μ^* 为由 μ 引出的外测度, $A\subset\Omega$, 若 $\mu^*(A)<\infty$ 且对任一 $\varepsilon>0$, 有一 $A_\varepsilon\in\mathscr{F}$ 存在, 使得 $\mu^*(A\,\triangle\,A_\varepsilon)<\varepsilon$, 则 A 为 μ^*-可测集. (提示: 取 $B=\bigcap\limits_{n=1}^{\infty}\bigcup\limits_{m=n}^{\infty}A_{\frac{1}{2^m}}$, 注意对任意正整数$k$, $B=\bigcap\limits_{n=k}^{\infty}\bigcup\limits_{m=n}^{\infty}A_{\frac{1}{2^m}}$, 证明$\mu^*(A-B)=\mu^*(B-A)=0$.因而$B,A-B,B-A$均为$\mu^*$- 可测集, 而$A=(A-B)\cup(B-(B-A))$.

6. 设 $(\Omega,\mathscr{A},P)$ 为概率空间, $\widetilde{\Omega}\subset\Omega$ 是外测度为 1 的集 (即 $\widetilde{\Omega}$ 具有性质: 如 $C\in\mathscr{A},\widetilde{\Omega}\subset C$, 则 $P(C)=1$), 则 $(\widetilde{\Omega},\widetilde{\mathscr{A}},\widetilde{P})$ 也是概率空间, 其中 $\widetilde{\mathscr{A}}=\mathscr{A}\cap\widetilde{\Omega},\widetilde{P}(B)=P(A)$, 如果 $B=A\cap\widetilde{\Omega}$.(注意: 需证 $\widetilde{P}$ 定义的合理性.)

7. 设 μ 是可测空间 $(\Omega,\mathscr{A})$ 上的测度, $(\Omega,\bar{\mathscr{A}},\bar{\mu})$ 是 $(\Omega,\mathscr{A},\mu)$ 的完全化, 若 $A,B\in\mathscr{A},A\subset E\subset B$ 且 $\mu(B-A)=0$, 则 $E\in\bar{\mathscr{A}}$.

8. 设 $\mathscr{S}$ 是半集代数, μ_1,μ_2是$\sigma(\mathscr{S})$ 上的测度, 且在 $\mathscr{S}$ 上 σ-有限. 若

$$\mu_1(A)\leqslant\mu_2(A),\quad A\in\mathscr{S},$$

则

$$\mu_1(A)\leqslant\mu_2(A),\quad A\in\sigma(\mathscr{S}).$$

若 μ_1,μ_2 在 $\sigma(\mathscr{S})$ 上 σ-有限, 结果如何?

9. 设 $\Omega=R,\mathscr{S}^{(1)}=\{[a,b):-\infty\leqslant a\leqslant b\leqslant+\infty,$当$a=-\infty$时$[a,b)=(-\infty,b)\}$, 若 μ 是 $\mathscr{S}^{(1)}$ 上的测度, 对一切 $a,b\in R,\mu([a,b))<\infty$ 试证: 存在不降左连续函数 $F(x),x\in R$ 使得 $\mu([a,b))=F(b)-F(a)$, 对一切 $a,b\in R$ 成立. 并且 $F(x)$ 在相差一个常数的意义上是唯一的.

10. 若 $F(x)$ 是 $R^{(1)}$ 上非降左连续函数, 则在 $\mathscr{B}^{(1)}$ 上存在一个测度 μ, 使得对一切 $a,b\in R$, 有

$$\mu([a,b))=F(b)-F(a).$$

§1.6　独立事件类

独立性是概率论中十分重要的概念, 它在实际应用中也很重要. 但在初等概率论中有些结果难以严格地证明. 例如, 两个独立随机变量的连续函数仍然独立, 以及独立随机变量的乘积的数学期望等于数学期望的乘积等命题 (这些命题将在第 2、3 章中分别予以证明). 在讨论这些问题时将用到独立事件类的概念及其扩张定理. 本节阐述独立类的概念并证明独立类的扩张定理, 一方面作为本章前述理论的应用, 一方面为今后的应用打下一定的基础.

1.6.1　独立事件类的定义

我们约定, 本节涉及到的一切事件 (集合) 都是在概率空间 $(\Omega,\mathscr{A},P)$ 之上的, 即 $A\in\mathscr{A}$, 且其概率由 $P(A)$ 给定.

定义 1　如果

$$P(A\cap B)=P(A)\cdot P(B),$$

则称事件 A 与 B 是独立的.

定义 2　$A_1,\cdots,A_n$ 是 n 个事件, 若对任意 $1\leqslant m\leqslant n$, 任意 $1\leqslant k_1<k_2<\cdots<k_m\leqslant n$ 都有

$$P\Big(\bigcap_{j=1}^{m}A_{k_j}\Big)=\prod_{j=1}^{m}P(A_{k_j}),$$

则称 $A_1,\cdots,A_n$ 是独立的.

显然, 若 $A_1,\cdots,A_n$ 独立, 必有 $A_1,\cdots,A_n$ 两两独立 (这时取 $m=2$), 反之, 若 $A_1,\cdots,A_n$ 两两独立, 则未必有 $A_1,\cdots,A_n$ 独立. 读者不难构造例子说明.

以下我们用 T 表示一个任意的指标集.

定义 3 设 $A_t, t\in T$ 是 $\mathscr{A}$ 中的事件类, 若对任意 T 的有限子集 $\{t_1,\cdots,t_n\}$,

$$P(A_{t_1}\cap\cdots\cap A_{t_n})=P(A_{t_1})\cdots P(A_{t_n}), \tag{1}$$

则称 $A_t, t\in T$ **独立.**

若对任意 $t\in T, \mathfrak{E}_t$ 表示 $\mathscr{A}$ 中的非空事件类, 且从每一 $\mathfrak{E}_t$ 中任意取出一事件 A_t 后, 事件 $A_t, t\in T$ 独立, 则称**事件类** $\mathfrak{E}_t, t\in T$ **独立.**

1.6.2 独立类扩张定理

我们来证明本节的主要定理.

定理 1 设 $(\Omega,\mathscr{A},P)$ 是一概率场, 若 $\mathfrak{E}_t, t\in T$ 是其中的事件类, 且满足下列条件:

i) 事件类 $\mathfrak{E}_t, t\in T$ 独立;

ii) 每一 $\mathfrak{E}_t$ 对有限交封闭, 即若 $A_t^{(1)},\cdots,A_t^{(m)}\in\mathfrak{E}_t$, 则 $\bigcap\limits_{k=1}^{m}A_t^{(k)}\in\mathfrak{E}_t$, 则 $\sigma(\mathfrak{E}_t)$, $t\in T$ 是独立的事件类.

证 i) 首先由定义 3 易知 $\mathfrak{E}_t, t\in T$ 独立的充分必要条件是对 T 的任意有限子集 $\{t_1,\cdots,t_n\}, n\geqslant 1, \mathfrak{E}_{t_k}, k=1,\cdots,n$ 独立, 因而只需对 $T=\{1,\cdots,n\}$ 的情形来证明.

ii) 设 $T=\{1,\cdots,n\}$, 令

$$\mathcal{M}_1=\{A_1\in\sigma(\mathfrak{E}_1): 对任意A_j\in\mathfrak{E}_j, j=2,\cdots,n, A_1,A_2,\cdots,A_n独立\}$$

往证 $\mathcal{M}_1=\sigma(\mathfrak{E}_1)$. 由于 $\mathcal{M}_1$ 包含 π-系 $\mathfrak{E}_1$, 故只需证 $\mathcal{M}_1$, 为 λ-系.

首先证 $\Omega\in\mathcal{M}_1$: 即证对任意 $A_j, j=2,\cdots,n, \Omega,A_2,\cdots,A_n$ 独立. 事实上, 任取 $1\leqslant k_1<k_2<\cdots<k_m\leqslant n$, 若 $k_1=1$, 则由 $A_2,\cdots,A_n$ 独立及 $P(\Omega)=1$ 知

$$\begin{aligned}P(\Omega\cap A_{k_2}\cap\cdots\cap A_{k_m})&=P(A_{k_2}\cap\cdots\cap A_{k_m})\\&=P(A_{k_2})\cdots P(A_{k_m})\\&=P(\Omega)P(A_{k_2})\cdots P(A_{k_m}).\end{aligned}$$

若 $k_1>1$, 知 $\{A_{k_1},\cdots,A_{k_m}\}\subset\{A_2,\cdots,A_n\}$, 再由 $A_2,\cdots,A_n$ 独立有:

$$P\left(\bigcap_{j=1}^{m}A_{k_j}\right)=\prod_{j=1}^{m}P(A_{k_j}).$$

这就证明了, $\Omega, A_2, \cdots, A_n$ 独立, 即 $\Omega \in \mathcal{M}_1$.

再证 $\mathcal{M}_1$ 对真差封闭: 若 $A, B \in \mathcal{M}_1$, 且 $A \subset B$, 往证对任何 $A_j \in \mathfrak{C}_j, j = 2, \cdots, n, B - A, A_2, \cdots, A_n$ 独立. 利用 n 个事件独立的定义及概率的减法性质可得: 当 $1 = k_1 < k_2 < \cdots < k_m \leqslant n$ 时, 由于 $A, B \in \mathcal{M}_1$, 我们有

$$\begin{aligned} P\left((B-A)\cap\left(\bigcap_{j=2}^{m} A_{k_j}\right)\right) &= P\left(B\cap\left(\bigcap_{j=2}^{m} A_{k_j}\right)\right) - P\left(A\cap\left(\bigcap_{j=2}^{m} A_{k_j}\right)\right) \\ &= P(B)\cdot\prod_{j=2}^{m} P(A_{k_j}) - P(A)\cdot\prod_{j=2}^{m} P(A_{k_j}) \\ &= (P(B)-P(A))\prod_{j=2}^{m} P(A_{k_j}) \\ &= P(B-A)\prod_{j=2}^{m} P(A_{k_j}); \end{aligned}$$

当 $1 < k_1 < \cdots < k_m \leqslant n$时, 由$A_2, \cdots, A_n$ 独立易知

$$P(A_{k_1}\cap\cdots\cap A_{k_m}) = \prod_{j=1}^{m} P(A_{k_j}).$$

这就证明了 $B - A, A_2, \cdots, A_n$ 独立. 最后证 $\mathcal{M}_1$ 对单调增序列的并封闭. 即若 $B_n \uparrow B, B_l \in \mathcal{M}_1, l = 1, 2, \cdots$, 则 $B \in \mathcal{M}_1$. 同样当取 $1 = k_1 < k_2 < \cdots < k_m \leqslant n$ 时, 由于概率的连续性及 $B_l \in \mathcal{M}_1, l = 1, 2, \cdots$, 我们有

$$\begin{aligned} P\left(B\cap\left(\bigcap_{j=2}^{m} A_{k_j}\right)\right) &= \lim_{l\to\infty} P\left(B_l\cap\left(\bigcap_{j=2}^{m} A_{k_j}\right)\right) \\ &= \lim_{l\to\infty} P(B_l)\prod_{j=2}^{m} P(A_{k_j}) \\ &= P(B)\prod_{j=2}^{m} P(A_{k_j}), \end{aligned}$$

当取 $1 < k_1 < \cdots < k_m \leqslant n$时, 由$A_2, \cdots, A_n$ 独立知

$$P\left(\bigcap_{j=1}^{m} A_{k_j}\right) = \prod_{j=1}^{m} P(A_{k_j}).$$

这就证明了 $\sigma(\mathfrak{C}_1), \mathfrak{C}_2, \cdots, \mathfrak{C}_n$独立, 且每一个类都是$\pi$-系.

iii) 令 $\mathcal{M}_2 = \{A_2 \in \sigma(\mathfrak{C}_2) : A_1 \in \sigma(\mathfrak{C}_1), A_j \in \mathfrak{C}_j, j = 3, \cdots, n\}$, 与 ii) 同理可证 $\sigma(\mathfrak{C}_1), \sigma(\mathfrak{C}_2), \mathfrak{C}_3, \cdots, \mathfrak{C}_n$ 独立.

再以同样手法进行 $n-2$ 次即可证明 $\sigma(\mathfrak{E}_k), k=1,\cdots,n$ 独立. 定理证毕.

推论 若事件 $A_t, t\in T$ 是独立的, 则事件类 $\{\varnothing, A_t, A_t^c, \Omega\}, t\in T$ 独立, 因而对任意 $t_1,\cdots,t_n\in T$

$$P\left(\bigcap_{k=1}^{n} A'_{t_k}\right)=\prod_{k=1}^{n} P(A'_{t_k}),$$

其中 $A'_{t_k}=A_{t_k}$或$A^c_{t_k}$.

证 因为 $\{A_t\}$ 上的最小 σ-代数是 $\{\varnothing, A_t, A_t^c, \Omega\}$.

这个推论可用较定理 1 简单得多的方法来证明, 我们证明定理 1 的目的不在于此, 而是为了以后的应用.

下面先举几例说明定理 1 的应用.

例 1 设有 n 个独立事件 $A_1,\cdots,A_n$, 则

$$P\left(\bigcup_{k=1}^{n} A_k\right)=1-\prod_{k=1}^{n}(1-P(A_k)). \tag{2}$$

这是因为

$$\bigcup_{k=1}^{n} A_k=\left(\bigcap_{k=1}^{n} A_k^c\right)^c.$$

故

$$P\left(\bigcup_{k=1}^{n} A_k\right)=1-P\left(\bigcap_{k=1}^{n} A_k^c\right). \tag{3}$$

由于 $A_k, k=1,\cdots,n$ 是独立的, 故由推论知 $P\left(\bigcap\limits_{k=1}^{n} A_k^c\right)=\prod\limits_{k=1}^{n} P(A_k^c)=\prod\limits_{k=1}^{n}(1-P(A_k))$, 代入 (3) 即得 (2) 式.

公式 (2) 在确定概率时有很多应用, 例如

(1a) 设甲、乙两地间的无线电通讯中有 n 个中继站, 而每一中继站中断的概率是 p, 试求甲、乙两地通讯中断的概率 p_0.

设 A_k 表示 "第 k 个中继站中断" 这一事件, $k=1,\cdots,n$, 而只要有一中继站中断, 甲、乙两地的通讯即告中断, 故甲、乙两地通讯中断这一事件为 $\bigcup\limits_{k=1}^{n} A_k$. 另外, 通常两个中继站工作状况是没有多大关系的, 因此可以设为 $A_1,\cdots,A_n$ 是独立的, 故由 (2) 式知

$$p_0=P\left(\bigcup_{k=1}^{n} A_k\right)=1-\prod_{k=1}^{n}(1-P(A_k))=1-(1-p)^n. \tag{4}$$

(1b) 有 100 个民兵, 每一民兵一次射击命中敌机的概率是 0.05, 试求在 100 个民兵对敌机进行一次射击而敌机被击中一次以上的概率 p_0.

这个问题的解法与上例一样, 其结果为

$$p_0 = 1 - (1 - 0.05)^{100} \approx 0.994.$$

例 2　设 $A_1, \cdots, A_n$ 是 n 个独立事件, 且 $P(A_k) = p, k = 1, \cdots, n$, 令

$$B_m = \Sigma A_1' \cap \cdots \cap A_n', A_k' = A_k \text{或} A_k^c, \tag{5}$$

而和式是对 $A_1', \cdots, A_n'$ 中恰有 m 个 $A_k' = A_k$ 的一切事件求和的, 试求 B_m 的概率.

这个问题在通常教科书中采用以下更直观的说法: 将 $\{A_k, A_k^c\}$ 叫做一次试验, 而 $\{A_k, A_k^c\}, k = 1, \cdots n$ 叫做 n 次独立试验, 而 B_m 称为在 n 次独立试验中 A_k 发生 m 次. 这种说法是有其广泛的实际意义的, 下面举一实例: 某工厂产品中出现废品的概率为 0.005, 求在任意取来的 10000 件产品中恰有 40 件废品的概率. 在这个特例中每取一个产品可以看成是一次试验, 抽第 k 个产品而抽得废品以 A_k 表示, 抽得合格品以 A_k^c 表示, 因此 $\{A_k, A_k^c\}$ 恰好表示第 k 次试验的结果, 而且根据实际情况可以认为 $A_1, \cdots, A_{10000}$ 独立, 且 $P(A_k) = 0.005$, 所要求的概率是 $P(B_{40})$.

我们现在来求概率 $P(B_m)$, 由于 B_m 的和式表达式中各项是不交的 (即互不相容), 故

$$P(B_m) = \Sigma P(A_1' \cap \cdots \cap A_n'). \tag{6}$$

而由定理 1 的推论知 $A_1', \cdots, A_n'$ 是独立的, 故

$$P(A_1' \cap \cdots \cap A_n') = P(A_1') \cdots P(A_n').$$

但由假设知 $P(A_k) = p, P(A_k^c) = 1-p$, 而$A_1', \cdots, A_n'$ 中恰有 m 个 A_k, 因而 $P(A_1' \cap \cdots \cap A_n') = p^m(1-p)^{n-m}$, 又由假设知 (5) 中的项数是 n 个事件中取 m 个的组合数, 即为 $\begin{pmatrix} n \\ m \end{pmatrix}$, 故由 (6) 知

$$P(B_m) = \begin{pmatrix} n \\ m \end{pmatrix} p^m (1-p)^{n-m}. \tag{7}$$

有时为了简单, 将 $P(B_m)$ 记作 $b(m, n, p)$, 称为 n 次独立试验中 A_k 出现 m 次的概率, 或者用 μ 表示 n 次独立试验中 A_k 出现的次数, 则 $P(B_m)$ 也可记成 $P(\mu = m)$, 通常还用 q 表示 $1 - p$, 于是 (7) 式写成

$$b(m, n, p) = P(\mu = m) = \begin{pmatrix} n \\ m \end{pmatrix} p^m q^{n-m}. \tag{8}$$

例 3　设有一种对象 X 在一段时间内是随机地出现的 (我们称这种现象为流), 它的出现满足下列三个条件:

1° **平稳性** 对任何 $t>0$ 及非负整数 k, 在时间区间 $(a,a+t]$ 内 X 出现 k 次的概率对任何 $a\geqslant 0$ 都一样, 即 "在时间区间 $(a,a+t]$ 内 X 出现 k 次"(用 $A_k(a,t)$ 表示) 的概率只与 k,t 有关, 而与 a 无关, 这个概率记成 $P_k(t)$. 实际上, 还要假设在有限时间间隔 t 内 X 只出现有限次. 并且 X 出现一次或一次以上的概率大于零.

2° **无后效性** 对任意 $a,t,k,A_k(a,t)$ 的概率与时刻 a 以前的 X 出现的状况无关, 确切地说, 若 $(a_\lambda,a_\lambda+t_\lambda],\lambda\in\Lambda$ 是不相交的区间类, 则事件类 $\mathfrak{C}_\lambda\hat{=}\{A_k(a_\lambda,t_\lambda),k=0,1,2,\cdots\},\lambda\in\Lambda$ 是独立的.

3° **普通性** 设 $\psi(t)=\sum\limits_{k=2}^{\infty}P_k(t)$, 则

$$\psi(t)=o(t)(t\to 0)\quad 或\quad \lim_{t\to 0}\frac{\psi(t)}{t}=0.$$

这表示在同一刹那 X 出现两次或两次以上实际上是不可能的.

具有平稳性、无后效性和普通性的流为最简单流. 它是很多实际对象的数学抽象, 因而它有着广泛的应用. 下面仅举几例来说明:

(3a) X 表示某城市内电话用户向电话局的呼唤;

(3b) X 表示一块放射性元素单个原子的蜕变;

(3c) 电视的传送是将一幅画面分成若干个 (例如 25^2 个) 长方格, 而将这些格中的亮度分成若干级后依次按亮度的级数传送, 由于所传送的画面的多样性, 亮度的改变可以认为是随机的, 因此用 X 表示亮度的一次改变, 可认为它满足以上三个性质;

(3d) 当电子管接通时, 电流 (或电压) 是由电子管阴极向阳极飞出的电子而产生的. 虽然电源的电压一定, 但由于电子的众多以及其他各种原因, 每个电子从阴极飞出也是随机的, 因而电流也出现波动, 这种现象在无线电电子学中称为散弹效应. 用 X 表示有一个电子从阴极飞出, 则当电子管工作正常时 (即电子学中所谓的稳态状态), 可以认为 X 的流是最简单流.

下面我们来求出 $P_k(t)$ 的表示式. 首先求 $P_0(t)$.

i) 由于 $A_0(0,lt)=\bigcap\limits_{m=0}^{l-1}A_0(mt,t),t>0,$ 及 $A_0(mt,t),m=0,1,\cdots,l-1$ 独立 (无后效性), 故

$$P_0(lt)=P(A_0(0,lt))=\prod_{m=0}^{l-1}P(A_0(mt,t))=[P_0(t)]^l. \tag{9}$$

因而 $P_0(1)=\left[P_0\left(\frac{1}{n}\right)\right]^n$, n 为任意正整数, 令

$$P_0(1)=\theta, \tag{10}$$

则 $P_0\left(\frac{1}{n}\right)=\theta^{\frac{1}{n}}$, 由 (9) 即知

$$P_0\left(\frac{k}{n}\right)=\theta^{\frac{k}{n}}, \qquad k=1,2,\cdots. \tag{11}$$

应用 (11) 来证明：对任意正数 t 都有

$$P_0(t)=\theta^t. \tag{12}$$

由于当 $0<t_1<t_2$ 时,

$$\begin{aligned}P_0(t_2)&=P(A_0(0,t_2))=P(A_0(0,t_1)\cap A_0(t_1,t_2-t_1))\\&=P(A_0(0,t_1)P(A_0(t_1,t_2-t_1))\leqslant P(A_0(0,t_1))\\&=P_0(t_1),\end{aligned}$$

即 $P_0(t)$ 是不增函数. 对每一正整数 n 有一正整数 l_n 存在使得

$$\frac{l_n-1}{n}<t\leqslant\frac{l_n}{n}, \tag{13}$$

因而 $\lim\limits_{n\to\infty}\frac{l_n}{n}=t$. 这就有

$$\theta^{\frac{l_n-1}{n}}=P_0\left(\frac{l_n-1}{n}\right)\geqslant P_0(t)\geqslant P_0\left(\frac{l_n}{n}\right)=\theta^{\frac{l_n}{n}}.$$

令 $n\to\infty$, 即得 (12).

最后我们来证明

$$0<\theta<1. \tag{14}$$

由于 $\theta=P_0(1)$, 故应有 $0\leqslant\theta\leqslant1$, 若 $\theta=0$, 则对任一给定的 $t>0$, 由 (12) 知对任何正整数 n 都有 $P_0\left(\frac{t}{n}\right)=0$, 因而

$$P\left(\sum_{k=1}^{\infty}A_k\left(a,\frac{t}{n}\right)\right)=\sum_{k=1}^{\infty}P_k\left(\frac{t}{n}\right)=1-P_0\left(\frac{t}{n}\right)=1.$$

$\left(\text{因为在有限的时间间隔内 } X \text{ 不能出现无限次, 因而} \sum\limits_{k=0}^{\infty}P_k\left(\frac{t}{n}\right)=1\right)$, 由于无后效性及定理 1 知事件 $\bigcup\limits_{k=1}^{\infty}A_k\left(\frac{lt}{n},\frac{t}{n}\right), l=0,1,\cdots,n-1$ 是独立的, 且 $\bigcap\limits_{l=0}^{n-1}\left[\bigcup\limits_{k=1}^{\infty}A_k\left(\frac{lt}{n},\frac{t}{n}\right)\right]\subset\sum\limits_{k=n}^{\infty}A_k(0,t)$, 故

$$\sum_{k=n}^{\infty}P_k(t)=P\left(\sum_{k=n}^{\infty}A_k(0,t)\right)\geqslant\prod_{l=0}^{n-1}P\left(\bigcup_{k=1}^{\infty}A_k\left(\frac{lt}{n},\frac{t}{n}\right)\right)=1$$

对任何正整数 n 都成立, 令 $n\to\infty$, 即得矛盾. 若 $\theta=1$, 则由 (12) 知对任何正数 t 都有 $P_0(t)=1$, 于是

$$P\left(\sum_{k=1}^{\infty}A_k(0,t)\right)=\sum_{k=1}^{\infty}P_k(t)=1-P_0(t)=0,$$

即 X 在 $(0,t]$ 中出现一次及一次以上的概率是零, 此亦与假设矛盾, 故 (14) 式成立. 由 (12) 及 (14) 即得

$$P_0(t)=e^{-\lambda t},\lambda\text{为一正常数}. \tag{15}$$

这就是我们第一步所要求的结论.

ii) 现在来求 $P_k(t)(k\geqslant 1)$ 的公式, 设 t 是一给定的正数, 并令

$$(0,t]=\left(0,\frac{t}{n}\right]+\cdots+\left(\frac{n-1}{n}t,t\right],$$
$$B_1\triangleq A_k(0,t)\cap H_1,$$
$$B_2\triangleq A_k(0,t)\cap H_2,$$

其中

$H_1\triangleq$"在每一子区间中 X 出现的次数不超过 1",

$H_2\triangleq$"至少一子区间中 X 出现的次数大于 1".

则易见

$$P_k(t)=P(B_1)+P(B_2). \tag{16}$$

而 B_1 表示在 n 个子区间中恰有 k 个子区间 X 出现的次数是 1(其余的 $n-k$ 个子区间 X 出现的次数当然是零), 由于无后效性, 事件类 $\left\{A_0\left(\frac{lt}{n},\frac{t}{n}\right),A_1\left(\frac{lt}{n},\frac{t}{n}\right)\right\}$, $l=0,1,\cdots,n-1$ 是独立的, 应用与求 $b(k,n,p)$ 相同的方法, 立刻可以证明

$$P(B_1)=\begin{pmatrix}n\\k\end{pmatrix}\left[P_1\left(\frac{t}{n}\right)\right]^k\left[P_0\left(\frac{t}{n}\right)\right]^{n-k}. \tag{17}$$

由于 (15) 及普通性, 在 $n\to\infty\left(\frac{t}{n}\to 0\right)$ 及 k 为常数时, 我们有

$$\begin{aligned}\left[P_0\left(\frac{t}{n}\right)\right]^{n-k}&=e^{-\lambda\cdot\frac{t}{n}\cdot(n-k)}=e^{-\lambda t}e^{\lambda k\cdot\frac{t}{n}}=e^{-\lambda t}(1+o(1)),\\ \left[P_1\left(\frac{t}{n}\right)\right]^k&=\left[1-P_0\left(\frac{t}{n}\right)-\psi\left(\frac{t}{n}\right)\right]^k\\ &=\left[1-e^{-\lambda\frac{t}{n}}+o\left(\frac{t}{n}\right)\right]^k\\ &=\left[\lambda\cdot\frac{t}{n}+o\left(\frac{t}{n}\right)\right]^k=\frac{(\lambda t)^k}{n^k}[1+o(1)].\end{aligned}$$

将此二式代入 (17) 式, 当 $n \to \infty$ 时, 即得

$$\begin{aligned}P(B_1) &= e^{-\lambda t}\frac{(\lambda t)^k}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k}[1+o(1)] \\ &= \frac{(\lambda t)^k}{k!}e^{-\lambda t}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) \\ &\quad \cdots\left(1-\frac{k-1}{n}\right)[1+o(1)] \to \frac{(\lambda t)^k}{k!}e^{-\lambda t}.\end{aligned} \tag{18}$$

另一方面, $P(B_2) \leqslant P(H_2)$, 而 $H_2 = \bigcup\limits_{l=0}^{n-1}\left(\sum\limits_{k=2}^{\infty} A_k\left(\frac{lt}{n}, \frac{t}{n}\right)\right)$, 故由普通性知

$$\begin{aligned}P(B_2) \leqslant P(H_2) &\leqslant \sum_{l=0}^{n-1} P\left(\sum_{k=2}^{\infty} A_k\left(\frac{lt}{n}, \frac{t}{n}\right)\right) = n\cdot\psi\left(\frac{t}{n}\right) \\ &= t\frac{\psi\left(\frac{t}{n}\right)}{\frac{t}{n}} \to 0(n\to\infty).\end{aligned} \tag{19}$$

将 (18), (19) 代入 (16), 并令 $n \to \infty$ 即得

$$P_k(t) = \lim_{n\to\infty}[P(B_1)+P(B_2)] = \frac{(\lambda t)^k}{k!}e^{-\lambda t}, \quad k \geqslant 1.$$

而当 $k=0$ 时, 上式即为 (15) 式. 故由 i), ii) 得出下列一般表达式

$$P_k(t) = \frac{(\lambda t)^k}{k!}e^{-\lambda t}, \qquad k = 0,1,2,\cdots. \tag{20}$$

(20) 就是我们要求的表达式, 其中只有一个待定的参数 λ, 它的实际意义及对于具体问题的求法将在以后建立起来的理论的基础上给出. 此处不再讨论了.

习题及补充

1. 若对事件 $A_1,\cdots,A_n$ 有 $P\left(\bigcap\limits_{k=1}^{n} A_k\right) = \prod\limits_{k=1}^{n} P(A_k)$, 是否能得出 $A_1,\cdots,A_n$ 独立? 如果能, 证明之, 如果不能, 举出反例.

2. $(\Omega,\mathscr{A},P)$ 是概率场, 若 $\{A_n\} \subset \mathscr{A}$ 且 $\sum\limits_{n=1}^{\infty} P(A_n) < \infty$, 则 $P(A_n\text{i.o.}) = 0\left(A_n\text{i.o.}\right.$ 表示有无穷多个 A_n 发生, 即 $\left.A_n\text{i.o.} = \bigcap\limits_{n=1}^{\infty}\bigcup\limits_{k\geqslant n} A_k\right)$, 若 $A_n, n=1,2,\cdots$ 独立, 且 $\sum\limits_{n=1}^{\infty} P(A_n) = \infty$, 则 $P(A_n\text{i.o.}) = 1$.

此命题称为 Borel 0–1 判据.

(提示：证第二结论时应用概率的上、下连续性及 $A_n, n = 1, 2, \cdots$ 的独立性，求得 $P(A_n \text{i.o.}) = 1 - \lim\limits_{n\to\infty} \lim\limits_{m\to\infty} \prod\limits_{k=n}^{m} P(A_k^c)$, 再利用熟知不等式 $e^{-x} \geqslant 1 - x, x \geqslant 0$, 易得结论. 此结果称为 Borel 0−1 判据, 见VII.1)

3. 设 $A_1, \cdots, A_n$ 独立而 $P(A_k) = p_k$, 试求：

i) 所有事件全不发生的概率;

ii) 诸事件中至少发生其一的概率;

iii) 恰好发生其一 (任何一件) 的概率, 并指出其理论根据.

4. 设事件类 $\mathscr{C}_t, t \in T$ 独立, 则向每一个 $\mathscr{C}_t$ 中增添下列各种事件时独立性仍能保持：

i) 概率为零的事件及概率为 1 的事件;

ii) 其中事件的真差; 特别是它们的余集;

iii) 其中事件的可数不交和;

iv) 其中事件序列的极限.

5. 若事件类 $\mathfrak{E}_t, t \in T$ 独立, 每一 $\mathfrak{E}_t$ 为 π-系且 $T = \sum\limits_{k=1}^{n} T_k$, 令

$$\mathscr{A}_{T_k} = \sigma\left\{ \bigcap_{j=1}^{m} A_{t_j}, A_{t_j} \in \mathfrak{E}_{t_j}, t_j \in T_k, m \geqslant 1 \right\},$$

则 $\mathscr{A}_{T_k}, k = 1, \cdots, n$ 独立.

第2章　随机变量与可测函数、分布函数与 Lebesgue-Stieltjes 测度

第 1 章我们建立了概率场 $(\Omega, \mathscr{A}, P)$, 从原则上说, 已经完全建立了概率论的逻辑基础. 但是还需要在此基础上给出一些基本概念, 首先是随机变量及其分布函数的概念.

随机变量是一般测度论中可测函数的特殊情形, 分布函数与 Lebesgue-Stieltjes (以下简记为 L-S) 测度对应. 本章将在直观基础上给出随机变量及其分布函数的定义, 而后再更广泛地讨论可测函数及一般分布函数与 L-S 测度的关系.

§2.1　随机变量及其分布函数的直观背景

例 1　在 §1.2 例 2 中, 我们建立了基本事件系 $\Omega = \{A_k : k \in Z_+\}, A_k$ 表示在 $(a, a+t]$ 内收到用户 k 次呼叫这一基本事件. 如果我们用 ξ 表示该公用电话系统在 $(a, a+t]$ 内收到用户呼叫的次数, 则 ξ 是取值在 Z_+ 中的变量, 当 $\{A_k\}$ 发生时, ξ 取值为 k, 它与通常数学分析中的函数相同之处在于值域是某一数集, 不同之处在于定义域不一定是数集, 而是基本事件空间 Ω, 即

$$\xi(A_k) = k, \qquad A_k \in \Omega. \tag{1}$$

由于在一次试验中究竟发生哪一个 $\{A_k\}$ 是随机的, 从而 ξ 的取值也具有随机性. 粗略地说, 我们称这种取值具有随机性的变量为随机变量. 这就是随机变量这种术语的由来, 也就是它的实际背景所在.

现在进一步分析上述变量 ξ 的特性. 将随机变量 $\xi = k$ 与随机事件 $\{A_k\}$ 相联系, 而 Ω 的一切子集作成一个 σ-代数 $\mathscr{A}$, 由 §1.6 例 5 可知

$$P(\{A_k\}) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \qquad k \in Z_+, \tag{2}$$

若 $A \in \mathscr{A}$ (即 $A \subset \Omega$), 则令

$$P(A) = \sum_{A_k \in A} P(\{A_k\}). \tag{3}$$

于是容易验证 (2), (3) 定义了 $\mathscr{A}$ 上的一个概率, 这样就对该电话系统在 $(a, a+t]$ 内收到用户呼叫的这种随机现象规定了一个概率场 $(\Omega, \mathscr{A}, P)$. ξ 是定义在 Ω 上的

函数. 除此而外, 它还应满足其他性质. 例如, ξ 取某一非负整数值 k 是一个随机事件, 更一般些, ξ 取区间 $[x_1, x_2)$ 内的值, 也是随机事件, 它可以表成 $\{x_1 \leqslant \xi < x_2\}$, 这种随机事件应与概率场发生关系, 由 (1) 易知, 随机事件 $\{x_1 \leqslant \xi < x_2\}$ 和随机事件 $\{A_k : x_1 \leqslant \xi(A_k) < x_2\}$ 是一样的. $\{x_1 \leqslant \xi < x_2\}$ 显然是随机事件 $\{\xi < x_2\}$ 与 $\{\xi < x_1\}$ 的差, 故我们只需考虑 $\{\xi < x\}$ (x 是任意实数) 就可以了. 于是就得知 ξ 应具有第二个性质: 对任何实数来说, 使得 ξ 取值小于 x 的一切基本事件所成的集合应属于 $\mathscr{A}$, 即

$$\{A_k : \xi(A_k) < x\} \in \mathscr{A}, x \in R. \tag{4}$$

由 (4) 我们还可以证明 (见 §2.1), 对任何实直线上的 Borel 集 $B, \{A_k : \xi(A_k) \in B\} \in \mathscr{A}$. 这就使得对 $\{A_k : \xi(A_k) \in B\}$ 这样的事件能够赋予概率, 从而通过概率场, 完全决定了随机变量 ξ 的取值规律.

在很多场合, 对于一种随机现象, 只引进一个随机变量是不够的, 我们举例来说明这一点.

例 2 某射手向一目标射击 (参看 §1.2 例 6), 弹着点散布在目标周围, 设想在平面上建立了直角坐标系, 则 "弹着点为 (x, y)" 是一基本事件, 记作 $\lambda_{(x,y)}$, 基本事件空间 $\Omega \triangleq \{\lambda_{(x,y)} : (x, y) \in R^{(2)}\}$, 令 σ-代数 $\mathscr{A}$ 为一切形如

$$\Lambda(B) = \{\lambda_{(x,y)} : (x, y) \in B\}, B \in \mathscr{B}^{(2)} \tag{5}$$

的集类, 令

$$P(\Lambda(B)) = \iint\limits_{B} \frac{1}{2_\pi} \mathrm{e}^{-\frac{x^2+y^2}{2}} \mathrm{d}x\mathrm{d}y, B \in \mathscr{B}^{(2)}, \tag{6}$$

($P(\Lambda(B))$ 是概率, 其中积分为 Lebesgue 积分, 这点将在以后说明). 这样我们就建立了一个概率场 $(\Omega, \mathscr{A}, P)$.

为了研究它, 我们常常引进两个随机变量, 用 ξ 和 η 分别表示弹着点在 X 轴和 Y 轴上的坐标. 它们取值的范围都是实数集, 并且对于每一给定的基本事件 $\lambda_{(x,y)}, \xi$ 取值 x, η 取值 y, 即 ξ, η 都是定义在基本事件空间 Ω 上的函数:

$$\begin{aligned} \xi(\lambda_{(x,y)}) &= x, \\ \eta(\lambda_{(x,y)}) &= y, \end{aligned} \qquad \lambda_{(x,y)} \in \Omega.$$

同例 1 一样, ξ, η 具有下述性质:

$$\{\lambda_{(x',y')} : \xi(\lambda_{(x',y')}) < x\} \in \mathscr{A}, 对一切 x \in R,$$

$$\{\lambda_{(x',y')} : \eta(\lambda_{(x',y')}) < y\} \in \mathscr{A}, 对一切 y \in R.$$

有时, 还需要引进复随机变量 (即取值在复平面上).

定义 1 设 $(\Omega, \mathscr{A}, P)$ 是一概率场, $\xi = \xi(\omega), \omega \in \Omega$, 是定义在 Ω 上的实值函数, 且对任何实数 x

$$\{\omega : \xi(\omega) < x\} \in \mathscr{A}. \tag{7}$$

则称 ξ 为 $(\Omega, \mathscr{A}, P)$ 上的一个**实随机变量**, 若 $\xi = \eta + i\zeta$, 而 η, ζ 是 $(\Omega, \mathscr{A}, P)$ 上的实随机变量, 则称 ξ 为 $(\Omega, \mathscr{A}, P)$ 上的一个**复随机变量**.

若 $\xi_1, \cdots, \xi_n$ 是 $(\Omega, \mathscr{A}, P)$ 上的 n 个实 (复) 随机变量, 则 Ω 上的向量函数 $(\xi_1, \cdots, \xi_n)$ 称为 $(\Omega, \mathscr{A}, P)$ 上的一个 **n 维实 (复) 随机变量或 n 维实 (复) 随机向量**.

由上述定义看出随机变量本身的描述只用到基本事件空间 Ω 和 σ-代数 $\mathscr{A}$, 因此, 在研究随机变量时, 可以暂不考虑概率, 而把随机变量看作定义在可测空间 $(\Omega, \mathscr{A})$ 上的可测函数的特例 (在 §2.1 详细讨论). 另一方面, 对一维随机变量来说, 实际上要求集合 $\{\omega : \xi(\omega) < x\}$ 应有概率存在. 当随机变量 ξ 给定时, 它的概率只与 x 有关, 而对 n 维实随机变量 $(\xi_1, \cdots, \xi_n)$, 当任给定实数 $x_1, \cdots, x_n$ 时,

$$\bigcap_{k=1}^{n} \{\omega : \xi_k(\omega) < x_k\} \in, \mathscr{A}, \tag{8}$$

即此集合是随机事件, 因而有概率存在, 且当 $(\xi_1, \cdots, \xi_n)$ 给定时, 它只与 $x_1, \cdots, x_n$ 有关, 故对给定的 $(x_1, \cdots, x_n)$, 概率

$$P\left(\bigcap_{k=1}^{n} \{\omega : \xi_k(\omega) < x_k\}\right), (x_1, \cdots, x_n) \in R^{(n)} \tag{9}$$

是一个定义在 n 维欧氏空间 $R^{(n)}$ 上的函数, 它可以作为研究随机变量的工具. 我们给出下列

定义 2 设 $(\xi_1, \cdots, \xi_n)(n \geqslant 1)$ 是 $(\Omega, \mathscr{A}, P)$ 上的一个 n 维实随机变量, 则由 (9) 确定的 $R^{(n)}$ 上的函数叫做 $(\xi_1, \cdots, \xi_n)$**的分布函数**, 记作 $F_{\xi_1, \cdots, \xi_n}(x_1, \cdots, x_n)$, 在不发生混淆的情况下，通常记作 $F(x_1, \cdots, x_n)$.

为了书写简单, 我们以后经常将 $\bigcap\limits_{k=1}^{n} \{\omega : \xi_k(\omega) < x_k\}$ 记作 $\bigcap\limits_{k=1}^{n} \{\xi_k < x_k\}$ 或 $\{\xi_1 < x_1, \cdots, \xi_n < x_n\}$. 我们经常采用表示式

$$F(x_1, \cdots, x_n) = P(\xi_1 < x_1, \cdots, \xi_n < x_n) = P\left(\bigcap_{k=1}^{n} \{\xi_k < x_k\}\right).$$

下面我们举几个例子, 有的例子只给出分布函数, 不指明它的概率场.

例 3 μ 表示 n 次独立试验中事件 A 出现的次数. 我们首先来构造概率场.(参看 §1.2 例 1): 用 ω_k 表示第 k 次试验 A 出现, ω_k^c 表示第 k 次试验 A 不出现, 则基本事件集合

$$\Omega=\{(\omega_1',\cdots,\omega_n'):\omega_k'=\omega_k\text{或}\omega_k^c,k=1,\cdots,n\}$$

由 2^n 个基本事件组成. 令 $\mathscr{A}$ 是 Ω 的一切子集作成的集系. 若 $\omega_1',\cdots,\omega_n'$ 中有 m 个 $\omega_k, n-m$ 个 ω_k^c, 则令

$$P(\{(\omega_1',\cdots,\omega_n')\})=p^m(1-p)^{n-m}.$$

对任何 $B\in\mathscr{A}$, 令

$$P(B)=\sum_{(\omega_1',\cdots,\omega_n')\in B}P(\{(\omega_1',\cdots,\omega_n')\}).$$

则易知 P 是 $\mathscr{A}$ 上的概率.

令

$$\mu((\omega_1',\cdots,\omega_n'))=m,\text{其中}\omega_1',\cdots,\omega_n'\text{中恰有}m\text{个}\omega_k.$$

则 μ 是定义在 Ω 上的函数. 因为 $\mathscr{A}$ 是 Ω 的一切子集作成的集系, 因而

$$\{(\omega_1',\cdots,\omega_n'):\mu((\omega_1',\cdots,\omega_n'))<x\}\in\mathscr{A}.$$

故 μ 是 $(\Omega,\mathscr{A},P)$ 上的随机变量, 且

$$\begin{aligned}P(\mu=m)&=P(\{(\omega_1',\cdots,\omega_n'):\omega_1',\cdots,\omega_n'\text{中恰有}m\text{个}\omega_k\})\\&=\begin{pmatrix}n\\m\end{pmatrix}p^m(1-p)^{n-m}.\end{aligned}$$

将此概率记作 $b(m,n,p), 1-p$ 记作 q, 则 μ 的分布函数是

$$F(x)=F_\mu(x)=\sum_{m<x}b(m,n,p)=\sum_{0\leqslant m<x}\begin{pmatrix}n\\m\end{pmatrix}p^mq^{n-m}. \tag{10}$$

此即熟知的**二项分布律**, 记作 $B(n,p)$, 而称随机变量 μ **为按二项分布律分布的**.

例 4 按多项分布律分布的 r 维随机变量. 考虑一个 n 次独立试验序列, 在第 k 次试验中可能出现的结果为 $A_k^{(1)},\cdots,A_k^{(r+1)}$, 且 $A_k^{(s)}$ 出现的可能性为 $p_s, p_1+p_2+\cdots+p_{r+1}=1$, 这是例 3 的推广.

现在我们对这个 n 次独立试验建立一个概率场. 令基本事件集

$$\Omega=\{(A_1^{(s_1)},\cdots,A_n^{(s_n)}):s_k=1,\cdots,r+1;k=1,\cdots,n\},$$

Ω 由 $(r+1)^n$ 个基本事件组成, $\mathscr{A}$ 是由 Ω 的一切子集作成的 σ-代数. 再令

$$P(\{(A_1^{(s_1)},\cdots,A_n^{(s_n)})\})=p_{s_1}\cdots p_{s_n}, \tag{11}$$

对一切 $A \in \mathscr{A}$, 定义

$$P(A) = \sum_{(A_1^{(s_1)},\cdots,A_n^{(s_n)})\in A} p_{s_1}\cdots p_{s_n}, \tag{12}$$

则 P 为 $\mathscr{A}$ 上的概率.

令 $\mu_s(s=1,\cdots,r)$ 表示 $(A_1^{(s_1)},\cdots,A_n^{(s_n)})$ 中 $A_k^{(s)}(k=1,\cdots,n)$ 的个数, 即

$$\mu_s((A_1^{(s_1)},\cdots,A_n^{(s_n)})) = m_s,$$

因而仿照例 3$(\mu_1,\cdots,\mu_r)$ 是 $(\Omega,\mathscr{A},P)$ 上的 r 维实随机变量.

下面我们来求 $(\mu,\cdots,\mu_r)$ 的分布函数. 首先由 (11), (12) 知对于满足条件 $m_1+\cdots+m_r \leqslant n$ 的非负整数 $m_1,\cdots,m_r$, 有

$$P(\mu_1=m_1,\cdots,\mu_r=m_r) = \sum_{\substack{\mu_s((A_1^{(s_1)},\cdots,A_n^{(s_n)}))=m_s\\ s=1,\cdots,r}} p_{s_1}\cdots p_{s_n}, \tag{13}$$

其中每一个被加项都有m_s个$p_s, s=1,\cdots,r$,因而有 $n-\sum\limits_{s=1}^{r} m_s$个$p_{r+1}$,即上式中每一被加项都是 $p_1^{m_1}\cdots p_r^{m_r}p_{r+1}^{n-\sum\limits_{s=1}^{r} m_s}$. 而由组合公式知上式共有 $\dfrac{n!}{m_1!\cdots m_r!\left(n-\sum\limits_{s=1}^{r} m_s\right)!}$ 项，故

$$P(\mu_1=m_1,\cdots,\mu_r=m_r) = \begin{cases} \dfrac{n!}{m_1!\cdots m_r!\left(n-\sum\limits_{s=1}^{r} m_s\right)!} p_1^{m_1} \\ \qquad \cdots p_r^{m_r}\cdot p_{r+1}^{\left(n-\sum\limits_{s=1}^{r} m_s\right)}, \\ \qquad m_s \geqslant 0 \\ \qquad m_1+\cdots+m_r \leqslant n, \\ 0, \quad \text{其他情形}. \end{cases}$$

因而

$$\begin{aligned} F(x_1,\cdots,x_r) &= F_{(\mu_1,\cdots,\mu_r)}(x_1,\cdots,x_r) \\ &= \sum_{\substack{m_s<x_s\\ s=1,\cdots,r}} P(\mu_1=m_1,\cdots,\mu_r=m_r). \end{aligned} \tag{14}$$

这个函数仿照二项分布律的命名可称为**多项分布律**.

例 5 Poisson(泊松) 分布 由例 1 的分析我们知道当 $t=1$ 时 ξ 的分布函数为

$$F(x)=F_\xi(x)=\sum_{k<x}\frac{\lambda^k}{k!}e^{-\lambda}, \tag{15}$$

这个函数称为具有参数 λ 的 Poisson**分布律**.而在时间间隔 $(a,a+t]$ 内电话局收到的用户呼唤次数 ξ 按具有参数 λt 的 Poisson 律分布.

Poisson 律还可以作为二项律的近似 (当 p 充分小而 n 充分大时) 计算公式. 这只要从下面的命题即可看出：若 $n\to\infty$, 有 $np_n\to\lambda(\lambda$ 为正常数), 则对每一给定的 m, 有

$$\lim_{n\to\infty}b(m,n,p_n)=\frac{\lambda^m}{m!}e^{-\lambda}. \tag{16}$$

事实上,

$$\begin{aligned}b(m,n,p_n)&=\binom{n}{m}p_n^m q_n^{n-m}=\frac{n!}{m!(n-m)!}p_n^m q_n^{n-m}\\&=\frac{n(n-1)\cdots(n-m+1)}{n^m}\frac{\lambda^m}{m!}\left(1-\frac{\lambda}{n}\right)^{n-m}.\end{aligned} \tag{17}$$

但当 m 一定, $n\to\infty$ 时,

$$\frac{n(n-1)\cdots(n-m+1)}{n^m}=\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{m-1}{n}\right)\to 1,$$

$$\left(1-\frac{\lambda}{n}\right)^m\to 1,\left(1-\frac{\lambda}{n}\right)^n=\left[\left(1-\frac{\lambda}{n}\right)^{-\frac{n}{\lambda}}\right]^{-\lambda}\to e^{-\lambda},$$

将这些式子代入 (17) 即得 (16).

例 6 多维 Poisson 分布 仿照 (16) 证明的方法不难证明：若对任何正整数 $n, np_{ns}=\lambda_s, s=1,\cdots,r$, 是 r 个正常数, 且 $p_{n1}+\cdots+p_{nr+1}=1, p_{ns}>0$, 则当 $m_1,\cdots,m_r$ 是给定的非负整数时,

$$\begin{aligned}&\lim_{n\to\infty}\frac{n!}{m_1!\cdots m_r!\left(n-\sum\limits_{s=1}^r m_s\right)!}p_{n1}^{m_1}\cdots p_{nr}^{m_r}p_{n(r+1)}^{\left(n-\sum\limits_{s=1}^r m_s\right)}\\&=\frac{\lambda_1^{m_1}\cdots\lambda_r^{m_r}}{m_1!\cdots m_r!}e^{-(\lambda_1+\cdots+\lambda_r)}.\end{aligned} \tag{18}$$

因此可以用 (18) 的右端来近似计算多项分布律.

设 $(\xi_1,\cdots,\xi_r)$ 是一 r 维随机变量,

$$P(\xi_1=m_1,\cdots,\xi_r=m_r)=e^{-(\lambda_1+\cdots+\lambda_r)}\frac{\lambda_1^{m_1}\cdots\lambda_r^{m_r}}{m_1!\cdots m_r!},$$

则 $(\xi_1,\cdots,\xi_r)$ 的分布函数是

$$F_{\xi_1,\cdots,\xi_r}(x_1,\cdots,x_r)=e^{-(\lambda_1+\cdots+\lambda_r)}\sum_{0\leqslant m_1<x_1}\cdots\sum_{0\leqslant m_r<x_r}\frac{\lambda_1^{m_1}\cdots\lambda_r^{m_r}}{m_1!\cdots m_r!},\tag{19}$$

其中 $\lambda_s, s=1,\cdots,r$ 是给定的 r 个正常数. 此时称函数 (19) 为具有参数 $\lambda_1,\cdots,\lambda_r$ 的 r 维 Poisson 分布律, 而 $(\xi_1,\cdots,\xi_r)$ 称为按 r 维 Poisson 律分布的随机变量.

例 7 均匀分布 设 $[a,b], a<b$ 是一有限区间, 属于它的点看成基本事件, $\Omega=[a,b], \mathscr{A}=\mathscr{B}^{(1)}\bigcap[a,b]$, 由 §1.3 习题 9 知其为 σ-代数. 定义

$$P(A)=\frac{\mu(A)}{b-a},\qquad A\in\mathscr{A},$$

其中 μ 是直线上的 Lebesgue 测度. 于是 P 是 $\mathscr{A}$ 上的概率.

设 ξ 是在 $[a,b]$ 上取值的随机变量, 且

$$\xi(y)=y,\quad y\in[a,b],$$

则 ξ 的分布函数是

$$F_\xi(x)=\begin{cases}0, & x\leqslant a,\\ \dfrac{x-a}{b-a}, & a<x<b,\\ 1, & x\geqslant b.\end{cases}\tag{20}$$

这个分布称为区间 $[a,b]$ 上的**均匀分布**. ξ 称为在 $[a,b]$ 上均匀分布的随机变量.

例 8 正态分布 设 ξ 是某一概率场上的实随机变量, 它的分布函数是

$$F_\xi(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-m)^2}{2\sigma^2}}\mathrm{d}t,\tag{21}$$

其中 $\sigma>0$, $m\in R$, 则称 ξ 是**按参数为 m,σ 的正态 (或Gauss) 律分布的**. 函数 (21) 称为具有参数 m,σ 的正态律 (Gauss 律), 记作 $N(m,\sigma^2)$.

例 9 多维正态律 设 $(\xi_1,\cdots,\xi_n)$ 是某一概率场上的 n 维实随机变量, 它的分布函数是

$$F(x_1,\cdots,x_n)=\frac{1}{(2\pi)^{\frac{n}{2}}|D|^{\frac{1}{2}}}\cdot$$

$$\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}\exp\left\{-\frac{1}{2}(t-m)D^{-1}(t-m)'\right\}\mathrm{d}t_1\cdots\mathrm{d}t_n,\tag{22}$$

其中 D 是 $n\times n$ 定正方阵 $(d_{ij}), D^{-1}$ 表示它的逆, $(t-m)$ 表示行向量 $(t_1-m_1,\cdots,t_n-m_n), (t-m)'$ 表示 $(t-m)$ 的转置, $|D|$ 表示方阵 D 的行列式. 则 $(\xi_1,\cdots,\xi_n)$ 称为**按 n 维正态律 (或 n 维 Gauss 律) 分布的.** 而函数 (22) 称为 n 维正态律 (或 Gauss 律), 它由向量 $\boldsymbol{m}$ 及方阵 D 决定, 因此记作 $N(\boldsymbol{m},D)$.

正态律在概率论中具有重要的地位, 因为很多重要生产现象及自然现象中的随机变量都可以认为是按正态律分布的, 例如很多工业产品的指标, 如棉纱的强力、支数, 螺丝钉的口径等, 都可以认为是按正态分布的. 若以 (ξ,η) 表示炮击一固定目标时弹着点的位置, 则 (ξ,η) 可认为是按二维正态分布的, 等等. 除此而外, 正态分布还可作为很多分布函数的近似 (例如当 n 充分大时, 在一定意义下二项分布便近似于正态分布), 因而可以大大地简化计算及促进问题的解决 (后一点可在数理统计中看到).

最后我们引进

定义 3 设 $(\xi_1,\cdots,\xi_n),(\eta_1,\cdots,\eta_n)$ 是概率场 $(\Omega,\mathscr{A},P)$ 上的两个 n 维随机变量, $n\geqslant 1$, 若

$$P\left(\bigcup_{k=1}^{n}\{\xi_k\neq\eta_k\}\right)=0,$$

则称 $(\xi_1,\cdots,\xi_n)$ 与 $(\eta_1,\cdots,\eta_n)$ 是**几乎相等**的. 若

$$F_{\xi_1,\cdots,\xi_n}(x_1,\cdots,x_n)=F_{\eta_1,\cdots,\eta_n}(x_1,\cdots,x_n),(x_1,\cdots,x_n)\in R,$$

则称 $(\xi_1,\cdots,\xi_n)$ 与 $(\eta_1,\cdots,\eta_n)$ 是**同分布**的.

今后我们将几乎相等的随机变量看成是相同的, 因为两个随机变量虽然作为 Ω 上的函数是不等的, 但是使它们不等的集合概率为零, 而按照直观理解, 它们不等的这个事件在实际上是不可能发生的, 因而这种约定是合理的.

容易证明, 几乎相等的随机变量是同分布的, 而同分布的随机变量未必是几乎相等的. 例如: ξ 按 $N(0,1)$ 分布, 则 $-\xi$ 也是按 $N(0,1)$ 分布的, 但 $P(\{\xi\neq-\xi\})=1$.

还容易证明: $(\xi_1,\cdots,\xi_n)$ 与 $(\eta_1,\cdots,\eta_n)$ 几乎相等的充分必要条件是对每一 $k(1\leqslant k\leqslant n),\xi_k$ 与 η_k 几乎相等.

习题与补充

1. 设 A 是概率场 $(\Omega,\mathscr{A},P)$ 上的一个随机事件, 试证

i) 对 A 可以定义一个只取 0, 1 为值的随机变量, 而它的分布函数

$$F_\xi(x)=\begin{cases}0, & x\leqslant 0,\\ P(A), & 0<x\leqslant 1,\\ 1, & x>1.\end{cases}$$

ii) 对 A 可以定义一个只取 -1 与 1 的随机变量 η, 而 η 的分布函数

$$F_\eta(x)=\begin{cases}0, & x\leqslant -1,\\ 1-P(A), & -1<x\leqslant 1,\\ 1, & x>1.\end{cases}$$

2. 设 r 是任一正整数, 在无穷独立试验序列中, 每次试验事件 A 出现的概率为 p, 用 ξ_r 表示第 ξ_r 次试验事件 A 恰恰第 r 次出现. 试求 ξ_r 的分布.(称 ξ_r 为**负二项分布**, 当 $r=1$ 时称为几何分布.)

3. 试证: 几乎相等的随机变量是同分布的, 而同分布的随机变量未必几乎相等.

4. $(\xi_1,\cdots,\xi_n)$ 与 $(\eta_1,\cdots,\eta_n)$ 几乎相等的充分必要条件是对每一 $k(1\leqslant k\leqslant n)$ 来说, ξ_k 与 η_k 几乎相等.

§2.2 随机变量与可测函数

随机变量的定义是取有限值的可测函数, 我们将会看到随机变量的极限则有可能取无限值, 因而是一般的可测函数, 为了从理论上更方便地研究随机变量, 本节将讨论一般的可测函数, 而把随机变量作为它的特例, 从而得出一些特殊性质.

可测函数可取 $\pm\infty$ 为值, 因而我们首先对 $\pm\infty$ 与有限数的运算规则作如下规定:

$$(\pm\infty)+x=x+(\pm\infty)=\pm\infty,\quad 当\quad |x|<\infty \tag{1}$$

$$(\pm\infty)+(\pm\infty)=\pm\infty, \tag{2}$$

$$x\cdot(\pm\infty)=(\pm\infty)\cdot x=\begin{cases}\pm\infty, & 若 0<x\leqslant+\infty,\\ 0, & 若 x=0,\\ \mp\infty, & 若 -\infty\leqslant x<0.\end{cases} \tag{3}$$

$$\frac{x}{\pm\infty}=0, 若 |x|<\infty. \tag{4}$$

我们还认为 $(\pm\infty)-(\pm\infty),(\pm\infty)+(\mp\infty),\dfrac{\pm\infty}{\pm\infty},\dfrac{x}{0}(x$ 任意) 是无意义的.

2.2.1 可测映射与可测函数

为了对实可测函数、复可测函数、n 维可测函数给出一个统一的概念和处理, 我们首先给出可测映射的定义.

定义 1 设 $(\Omega,\mathscr{A}),(\mathscr{X},\mathscr{B})$ 是两个可测空间, f 是从 Ω 到 $\mathscr{X}$ 的映射, 如果对一切 $B\in\mathscr{B}$,

$$\{\omega\in\Omega:f(\omega)\in B\}\in\mathscr{A},$$

则称 f 是由 $(\Omega,\mathscr{A})$ 到 $(\mathscr{X},\mathscr{B})$**的可测映射**.

定义 2 若 f 是从 $(\Omega,\mathscr{A})$ 到 $(\tilde{R}^{(1)},\tilde{\mathscr{B}}^{(1)})$ 的可测映射, 则称 f 为**实可测函数**; 若 f 是从 $(\Omega,\mathscr{A})$ 到 $(\tilde{R}^{(n)},\tilde{\mathscr{B}}^{(n)})$ 的可测映射, 则称 f 为 **n 维实可测函数**; 若 f 是从 $(\Omega,\mathscr{A})$ 到 $(\tilde{Z}^{(1)},\tilde{\mathscr{B}}_z^{(1)})$ 的可测映射, 则称 f 为**复可测函数**; 若 f 是从 $(\Omega,\mathscr{A})$ 到 $(\tilde{Z}^{(n)},\tilde{\mathscr{B}}_z^{(n)})$ 的可测映射, 则称 f 为 **n 维复可测函数**.

定义 3 设 f 是由 Ω 到 $\mathscr{X}$ 的映射, 对于 $\mathscr{X}$ 的任一子集 B, 称 Ω 的子集 $\{\omega: f(\omega)\in B\}$ 为 **B 在 f 之下的逆象**, 记作 $f^{-1}(B)$.

若 $\mathfrak{E}$ 是 $\mathscr{X}$ 的子集作成的一个类, 则 $\mathfrak{E}$ 中的一切集在 f 之下的逆象作成的类, 即 $\{f^{-1}(B): B\in\mathfrak{E}\}$, 称为 **$\mathfrak{E}$ 在 f 之下的逆象**, 记作 $f^{-1}(\mathfrak{E})$.

在讨论可测映射的性质之前, 先来讨论关于逆象的性质.

性质 1 设 f 是由 Ω 到 $\mathscr{X}$ 的映射, 则下列公式成立:

$$f^{-1}(\mathscr{X})=\Omega, f^{-1}(\varnothing)=\varnothing; \tag{5}$$

$$f^{-1}(B^c)=[f^{-1}(B)]^c, B\subset\mathscr{X}; \tag{6}$$

$$f^{-1}(B_1-B_2)=f^{-1}(B_1)-f^{-1}(B_2), B_1\subset\mathscr{X}, B_2\subset\mathscr{X}; \tag{7}$$

$$f^{-1}\left(\bigcup_{\gamma\in\Gamma}B_\gamma\right)=\bigcup_{\gamma\in\Gamma}f^{-1}(B_\gamma), B_\gamma\subset\mathscr{X}, \gamma\in\Gamma; \tag{8}$$

$$f^{-1}\left(\bigcap_{\gamma\in\Gamma}B_\gamma\right)=\bigcap_{\gamma\in\Gamma}f^{-1}(B_\gamma), B_\gamma\subset\mathscr{X}, \gamma\in\Gamma, \tag{9}$$

其中 Γ 是任一指标集.

证 由定义 3 易知 (5) 式成立. 再由定义 3 知, 若 $\omega\in f^{-1}(B^c)$, 则 $f(\omega)\notin B$, 即 $\omega\notin f^{-1}(B)$, 即 $\omega\in[f^{-1}(B)]^c$; 反之, 若 $\omega\in[f^{-1}(B)]^c$, 则 $\omega\notin f^{-1}(B)$, 因而 $f(\omega)\notin B$, 故 (6) 成立.

再来证明 (8), (9) 两式. 由定义 3 知

$$\begin{aligned}f^{-1}\left(\bigcup_{\gamma\in\Gamma}B_\gamma\right)&=\left\{\omega: f(\omega)\in\bigcup_{\gamma\in\Gamma}B_\gamma\right\}\\&=\{\omega: f(\omega)\text{至少属于某一}B_\gamma, \gamma\in\Gamma\}\\&=\bigcup_{\gamma\in\Gamma}\{\omega: f(\omega)\in B_\gamma\}=\bigcup_{\gamma\in\Gamma}f^{-1}(B_\gamma);\end{aligned}$$

$$\begin{aligned}f^{-1}\left(\bigcap_{\gamma\in\Gamma}B_\gamma\right)&=\left\{\omega: f(\omega)\in\bigcap_{\gamma\in\Gamma}B_\gamma\right\}\\&=\{\omega: f(\omega)\text{属于一切}B_\gamma, \gamma\in\Gamma\}\\&=\bigcap_{\gamma\in\Gamma}\{\omega: f(\omega)\in B_\gamma\}=\bigcap_{\gamma\in\Gamma}f^{-1}(B_\gamma).\end{aligned}$$

最后由 (6) 及 (9) 即得

$$\begin{aligned} f^{-1}(B_1 - B_2) &= f^{-1}(B_1 \cap B_2^c) = f^{-1}(B_1) \cap f^{-1}(B_2^c) \\ &= f^{-1}(B_1) \cap [f(B_2)]^c = f^{-1}(B_1) - f^{-1}(B_2). \end{aligned}$$

此即 (7) 式, 性质 1 获证. □

性质 2 设 $\mathscr{B}$ 是 $\mathscr{X}$ 中的 σ-代数, 则 $f^{-1}(\mathscr{B})$ 是 Ω 中的 σ-代数.

证 显然由 (5) 知

$$\Omega = f^{-1}(\mathscr{X}) \in f^{-1}(\mathscr{B});$$

再若 $A \in f^{-1}(\mathscr{B})$, 则有一 $B \in \mathscr{B}$, 使 $A = f^{-1}(B)$, 故知 $A^c = [f^{-1}(B)]^c = f^{-1}(B^c) \in f^{-1}(\mathscr{B})$, 即 $f^{-1}(\mathscr{B})$ 对求余运算封闭. 最后, 若 $\{A_n\}$ 是 $f^{-1}(\mathscr{B})$ 中的集序列, 则有 $\mathscr{B}$ 中的集序列 $\{B_n\}$ 存在, 使得 $A_n = f^{-1}(B_n), n = 1, 2, \cdots$, 于是由 (8) 知

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} f^{-1}(B_n) = f^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right) \in f^{-1}(\mathscr{B}).$$

故性质 2 获证. □

由性质 2 知, 若 f 是由 $(\Omega, \mathscr{A})$ 到 $(\mathscr{X}, \mathscr{B})$ 的可测映射, 令 $\sigma(f) \triangleq f^{-1}(\mathscr{B})$, 则 $\sigma(f)$ 是 σ-代数且 $\sigma(f) \subset \mathscr{A}$.

性质 3 若 $\mathfrak{C}$ 是 $\mathscr{X}$ 中的任一非空集类, 则 $\mathfrak{C}$ 上的最小 σ-代数的逆象为 $\mathfrak{C}$ 的逆象上的最小 σ-代数, 即

$$f^{-1}(\sigma(\mathfrak{C})) = \sigma(f^{-1}(\mathfrak{C})). \tag{10}$$

证 由性质 2 知 $f^{-1}(\sigma(\mathfrak{C}))$ 为一 σ-代数, 且显然有

$$f^{-1}(\sigma(\mathfrak{C})) \supset f^{-1}(\mathfrak{C}),$$

故有

$$f^{-1}(\sigma(\mathfrak{C})) \supset \sigma(f^{-1}(\mathfrak{C})).$$

今往证

$$f^{-1}(\sigma(\mathfrak{C})) \subset \sigma(f^{-1}(\mathfrak{C})). \tag{11}$$

我们令

$$\mathscr{C} = \{C : f^{-1}(C) \in \sigma(f^{-1}(\mathfrak{C}))\}, \tag{12}$$

于是

$$f^{-1}(\mathscr{C}) \subset \sigma(f^{-1}(\mathfrak{C})).$$

欲证 (11), 只需证 $\sigma(\mathfrak{E}) \subset \mathscr{C}$. 由于 $\sigma(\mathfrak{E})$ 是 $\mathfrak{E}$ 上最小 $\sigma-$ 代数, 故只需证 $\mathscr{C}$ 是包含 $\mathfrak{E}$ 的 $\sigma-$ 代数即可.

首先, 由定义 3 立知 $\mathfrak{E} \subset \mathscr{C}$; 其次由 (5) 知 $f^{-1}(\mathscr{X}) = \Omega \in \sigma(f^{-1}(\mathfrak{E}))$, 故由 $\mathscr{C}$ 的定义知 $\mathscr{X} \in \mathscr{C}$; 第三, 若 $C \in \mathscr{C}$, 则由 (6) 及 (12) 知 $f^{-1}(C^c) = [f^{-1}(C)]^c \in \sigma(f^{-1}(\mathfrak{E}))$, 因而 $C^c \in \mathscr{C}$; 最后, 若 $\{C_n\} \subset \mathscr{C}$, 则由 (8) 及 (12) 知 $f^{-1}(\bigcup\limits_n C_n) = \bigcup\limits_n f^{-1}(C_n) \in \sigma(f^{-1}(\mathfrak{E}))$, 因而 $\bigcup\limits_n C_n \in \mathscr{C}$, 故 $\mathscr{C}$ 为 $\sigma-$ 代数. 因而 $\sigma(\mathfrak{E}) \subset \mathscr{C}$, 即

$$f^{-1}(\sigma(\mathfrak{E})) \subset f^{-1}(\mathscr{C}) \subset \sigma(f^{-1}(\mathfrak{E})).$$

性质 3 获证. □

定理 1 i) f 是 $(\Omega, \mathscr{A})$ 上的实可测函数的充分必要条件是对一切 $x \in R^{(1)}$, $\{\omega : f(\omega) < x\} \in \mathscr{A}$;

ii) $f = (f_1, \cdots, f_n)$ 是 $(\Omega, \mathscr{A})$ 上的 n 维实可测函数的充分必要条件是对每一 $k = 1, \cdots, n$, f_k 是 $(\Omega, \mathscr{A})$ 上的实可测函数;

iii) $f = g + ih$ 是 $(\Omega, \mathscr{A})$ 上的复可测函数的充分必要条件是 g, h 是 $(\Omega, \mathscr{A})$ 上的实可测函数;

iv) $f = (f_1, \cdots, f_n)$ 是 $(\Omega, \mathscr{A})$ 上的 n 维复可测函数的充分必要条件是每一 f_k 是 $(\Omega, \mathscr{A})$ 上的复可测函数.

证 i) 条件的必要性显然, 今证充分性. 为此令

$$\mathfrak{E} = \{[-\infty, x), x \in R^{(1)}\}.$$

$\mathfrak{E}$ 是 $\tilde{R}^{(1)}$ 中的子集类, 易知 $\sigma(\mathfrak{E}) = \mathscr{B}^{(1)}$. 因此由 $\{\omega : f(\omega) < x\} \in \mathscr{A}, x \in R^{(1)}$ 成立, 知

$$f^{-1}(\mathfrak{E}) \subset \mathscr{A}.$$

故由 $\mathscr{A}$ 是 $\sigma-$ 代数有

$$\sigma(f^{-1}(\mathfrak{E})) \subset \mathscr{A}.$$

但由性质 3 知

$$f^{-1}(\tilde{\mathscr{B}}^{(1)}) = f^{-1}(\sigma(\mathfrak{E})) = \sigma(f^{-1}(\mathfrak{E})) \subset \mathscr{A}.$$

此即 f 是 $(\Omega, \mathscr{A})$ 上的实可测函数.

ii) 先证必要性: 若 $f = (f_1, \cdots, f_n)$ 是 $(\Omega, \mathscr{A})$ 上 n 维实可测函数, 今证每一 f_k 是 $(\Omega, \mathscr{A})$ 上的实可测函数. 为此, 只需证对任意的 $B_k \in \tilde{\mathscr{B}}^{(1)}, f_k^{-1}(B_k) \in \mathscr{A}$. 取

$B = \underbrace{\tilde{R}^{(1)} \times \cdots \times \tilde{R}^{(1)}}_{k-1\text{个}} \times B_k \times \underbrace{\tilde{R}^{(1)} \times \cdots \times \tilde{R}^{(1)}}_{(n-k)\text{个}}$, 则 $B \in \tilde{\mathscr{B}}^{(n)}$, 因而 $f^{-1}(B) \in \mathscr{A}$. 但

$$\begin{aligned} f^{-1}(B) &= \{\omega : (f_1(\omega), \cdots, f_n(\omega)) \in B\} \\ &= \bigcap_{i=1}^{k-1} \{\omega : f_i(\omega) \in \tilde{R}^{(1)}\} \cap \{\omega : f_k(\omega) \in B_k\} \\ &\quad \bigcap \left(\bigcap_{i=k+1}^{n} \{\omega : f_i(\omega) \in \tilde{R}^{(1)}\} \right) \\ &= \{\omega : f_k(\omega) \in B_k\} = f_k^{-1}(B_k), \end{aligned}$$

因而 $f_k^{-1}(B_k) \in \mathscr{A}$. 故每一 f_k 是 $(\Omega, \mathscr{A})$ 上可测函数.

再证充分性: 由 §1.3 例 11 对广义 n 维 Borel 域的推广知, 若令

$$\mathfrak{E} = \{B_1 \times \cdots \times B_n : B_k \in \tilde{\mathscr{B}}^{(1)}, k = 1, \cdots, n\},$$

则

$$\sigma(\mathfrak{E}) = \tilde{\mathscr{B}}^{(n)}.$$

故若 $f_1, \cdots, f_n$ 是 $(\Omega, \mathscr{A})$ 上实可测函数, 则

$$f^{-1}(\mathfrak{E}) \subset \mathscr{A}.$$

同 i) 一样可得

$$f^{-1}(\tilde{\mathscr{B}}^{(n)}) \subset \mathscr{A}.$$

因而 f 是 $(\Omega, \mathscr{A})$ 上 n 维实可测函数.

iii) 必要性: 若 $f = g + ih$ 是 $(\Omega, \mathscr{A})$ 上复可测函数, 今证 g 是 $(\Omega, \mathscr{A})$ 上实可测函数. 对任意的 $B_1 \in \tilde{\mathscr{B}}^{(1)}$, 取 $B = \{(x + iy) : x \in B_1, y \in \tilde{R}^{(1)}\}$, 则 $B \in \tilde{\mathscr{B}}_z^{(1)}$, 而由 f 是 $(\Omega, \mathscr{A})$ 到 $(\tilde{Z}^{(1)}, \tilde{\mathscr{B}}_z^{(1)})$ 上的可测映射知 $f^{-1}(B) \in \mathscr{A}$, 但

$$\begin{aligned} f^{-1}(B) &= \{\omega : g(\omega) + ih(\omega) \in B\} \\ &= \{\omega : g(\omega) \in B_1\} \cap \{\omega : h(\omega) \in \tilde{R}^{(1)}\} \\ &= g^{-1}(B_1) \cap \Omega = g^{-1}(B_1), \end{aligned}$$

故 $g^{-1}(B_1) \in \mathscr{A}$, 因而 g 是 $(\Omega, \mathscr{A})$ 上的实可测函数. 同理可证 h 也是 $(\Omega, \mathscr{A})$ 上实可测函数.

充分性: 设 $\mathfrak{E}$ 是一切下述形式的集所成的类:

$$B = \{(x + iy) : x \in B_1, y \in B_2\}, B_1, B_2 \in \tilde{\mathscr{B}}^{(1)},$$

则 $\sigma(\mathfrak{E})=\tilde{\mathscr{B}}_z^{(1)}$. 因此, 若 g,h 是 $(\Omega,\mathscr{A})$ 上实可测函数, 则易知

$$\begin{aligned}f^{-1}(B)=&\{\omega:g(\omega)+ih(\omega)\in B\}\\=&\{\omega:g(\omega)\in B_1\text{且}h(\omega)\in B_2\}\\=&\{\omega:g(\omega)\in B_1\}\cap\{\omega:h(\omega)\in B_2\}\in\mathscr{A}.\end{aligned}$$

故 $f^{-1}(\mathfrak{E})\subset\mathscr{A}$. 同 ii) 一样可得

$$f^{-1}(\tilde{\mathscr{B}}_z^{(1)})\subset\mathscr{A}.$$

即 f 是 $(\Omega,\mathscr{A})$ 上的复可测函数.

iv) 由于 n 维广义复 Borel 域 $\tilde{\mathscr{B}}_z^{(n)}$ 是 n 个一维广义复 Borel 域 $\tilde{\mathscr{B}}_z^{(1)}$ 的乘积 $\sigma-$ 代数, 故仿照 ii) 可证 iv). 不再仔细写出证明了. □

推论 $(\Omega,\mathscr{A})$ 上的一维实随机变量是 $(\Omega,\mathscr{A})$ 到 $(R^{(1)},\mathscr{B}^{(1)})$ 的可测映射, n 维实 (或复) 随机变量则是从 $(\Omega,\mathscr{A})$ 到 $(R^{(n)},\mathscr{B}^{(n)})$(相应地 $(Z^{(n)},\mathscr{B}_z^{(n)})$) 的可测映射.

应用定理 1 并注意到随机变量只取有限值, 即可知此推论成立.

由此推论知, 随机变量 (一维或 n 维, 实或复) 是取有限值的可测函数 (相应地, 一维或多维, 实或复).

为了进一步研究可测函数的性质, 我们给出复合映射的定义, 并证明

定理 2 若 $(\Omega_i,\mathscr{A}_i), i=1,2,3,$ 是三个可测空间, f 是从 $(\Omega_1,\mathscr{A}_1)$ 到 $(\Omega_2,\mathscr{A}_2)$ 的可测映射, g 是从 $(\Omega_2,\mathscr{A}_2)$ 到 $(\Omega_3,\mathscr{A}_3)$ 的可测映射, 则用

$$(g\circ f)(\omega_1)\triangleq g(f(\omega_1)),\quad \omega_1\in\Omega_1,$$

定义的**复合映射**$g\circ f$ 是 $(\Omega_1,\mathscr{A}_1)$ 到 $(\Omega_3,\mathscr{A}_3)$ 的可测映射.

证 首先, 由于 $f(\omega_1)\in\Omega_2$, 而 $g(f(\omega_1))\in\Omega_3$, 所以 $g\circ f$ 是 Ω_1 到 Ω_3 中的映射. 再者, 对任意的 $B\subset\Omega_3$ 有

$$\begin{aligned}(g\circ f)^{-1}(B)=&\{\omega_1:g(f(\omega_1))\in B\}\\=&\{\omega_1:f(\omega_1)\in g^{-1}(B)\}=f^{-1}(g^{-1}(B)).\end{aligned}$$

因此, 由 g 是 $(\Omega_2,\mathscr{A}_2)$ 到 $(\Omega_3,\mathscr{A}_3)$ 的可测映射知: 若 $B\in\mathscr{A}_3$, 则 $g^{-1}(B)\in\mathscr{A}_2$; 又由 f 是 $(\Omega_1,\mathscr{A}_1)$ 到 $(\Omega_2,\mathscr{A}_2)$ 的可测映射及 $g^{-1}(B)\in\mathscr{A}_2$ 知 $f^{-1}(g^{-1}(B))\in\mathscr{A}_1$. 故 $g\circ f$ 是 $(\Omega_1,\mathscr{A}_1)$ 到 $(\Omega_3,\mathscr{A}_3)$ 的可测映射. □

由定理 2 的特殊化立刻得到

定理 3 i) 设 g 是 $(\tilde{R}^{(n)},\tilde{\mathscr{B}}^{(n)})$ 上的实 (或复) 可测函数, 而 $f_1,\cdots,f_n$ 是可测空间 $(\Omega,\mathscr{A})$ 上的 n 个实可测函数, 则 $g(f_1,\cdots,f_n)$ 是 $(\Omega,\mathscr{A})$ 上的实 (或复) 可测函数.

ii) 设 g 是 $(\tilde{Z}^{(n)},\tilde{\mathscr{B}}_z^{(n)})$ 上的实 (或复) 可测函数, $f_1,\cdots,f_n$ 是 $(\Omega,\mathscr{A})$ 上的 n 个复可测函数, 则 $g(f_1,\cdots,f_n)$ 是 $(\Omega,\mathscr{A})$ 上实 (或复) 可测函数.

注 称满足定理条件的函数 g 为 Borel**可测函数**.

证 类似地, 只证 i). 令

$$f(\omega)=(f_1(\omega),\cdots,f_n(\omega)),\quad \omega\in\Omega.$$

则由定理 1 知 f 是 $(\Omega,\mathscr{A})$ 到 $(\tilde{R}^{(n)},\tilde{\mathscr{B}}^{(n)})$ 的可测映射, 而 g 是 $(\tilde{R}^{(n)},\tilde{\mathscr{B}}^{(n)})$ 到 $(\tilde{R}^{(1)},\tilde{\mathscr{B}}^{(1)})$(或 $(\tilde{Z}^{(1)},\tilde{\mathscr{B}}_z^{(1)})$) 的可测映射, 由定理 2 可知 $g\circ f$ 是 $(\Omega,\mathscr{A})$ 到 $(\tilde{R}^{(1)},\tilde{\mathscr{B}}^{(1)})$(相应地 $(\tilde{Z}^{(1)},\tilde{\mathscr{B}}_z^{(1)})$) 的可测映射, 即 $g(f_1,\cdots,f_n)$ 是 $(\Omega,\mathscr{A})$ 上的实 (或复) 可测函数. □

这个定理加以特殊化就成为

定理 4 i) 设 g 是 $(\tilde{R}^{(n)},\tilde{\mathscr{B}}^{(n)})$ 上的实 (或复) 可测函数, $f_1,\cdots,f_n$ 是 $(\Omega,\mathscr{A},P)$ 上的 n 个实随机变量. 若

$$P(A_\infty)=0,$$

其中

$$A_\infty=\{\omega:|g(f_1(\omega),\cdots,f_n(\omega))|=\infty\},\tag{13}$$

成立, 则 $g(f_1,\cdots,f_n)$ 是 $(\Omega,\mathscr{A},P)$ 上的实 (或复) 随机变量;

ii) 设 g 是 $(\tilde{Z}^{(n)},\tilde{\mathscr{B}}_z^{(n)})$ 上的实 (或复) 可测函数, $f_1,\cdots,f_n$ 是 $(\Omega,\mathscr{A},P)$ 上的复随机变量, 且 (13) 式成立, 则 $g(f_1,\cdots,f_n)$ 是 $(\Omega,\mathscr{A},P)$ 上的实 (复) 随机变量.

实际上, $g(f_1,\cdots,f_n)$ 与一个随机变量几乎相等.(即它与一个随机变量不等的事件概率为零).

这个定理在对随机变量进行理论讨论中是很重要的. 但是对随机变量的一个具体的函数, 例如 $\sin\xi, e^{it\xi}$ 等, 要想用这个定理断定是否为随机变量则还需要了解 Borel 可测函数及一般可测函数的构造性质. 因此我们在下一小节来讨论这些性质.

2.2.2 可测函数的构造性质

定义 4 i) 若 $A\subset\Omega$, 称函数 χ_A,

$$\chi_A(\omega)=\begin{cases}1, & 若\omega\in A,\\ 0, & 若\omega\notin A,\end{cases}$$

为集 A 的示性函数.

ii) 若 $A_k\in\mathscr{A},k=1,\cdots,n$, 两两不交且 $\sum\limits_{k=1}^{n}A_k=\Omega,a_1,\cdots,a_n$ 为实数、复数或 $\pm\infty$, 则称函数 f,

$$f(\omega)=\sum_{k=1}^{n}a_k\chi_{A_k}(\omega),\quad \omega\in\Omega,$$

为 $(\Omega,\mathscr{A})$**上的简单函数**.

iii) 若 $A_n\in\mathscr{A},n=1,2,\cdots$ 两两不交, 且 $\sum_{n=1}^{\infty}A_n=\Omega,a_n\in\tilde{Z}^{(1)},n=1,2,\cdots$, 则称函数 f,

$$f(\omega)=\sum_{n=1}^{\infty}a_n\chi_{A_n}(\omega),\omega\in\Omega,$$

为 $(\Omega,\mathscr{A})$ 上的**初等函数.**

性质 4 $(\Omega,\mathscr{A})$ 中可测集的示性函数、简单函数、初等函数都是 $(\Omega,\mathscr{A})$ 上可测函数.

证 只证初等函数的可测性, 其他两种情形是初等函数的特例. 设 $A_n\in\mathscr{A},n\geqslant 1$,

$$f(\omega)=\sum_{n=1}^{\infty}a_n\chi_{A_n}(\omega),\quad \omega\in\Omega.$$

对任一 $B\in\tilde{\mathscr{B}}^{(1)}$(或当 f 是复值函数时以 $\tilde{\mathscr{B}}_z^{(1)}$代替$\tilde{\mathscr{B}}^{(1)}$) 我们有

$$\begin{aligned}f^{-1}(B)&=\{\omega\colon f(\omega)\in B\}=\bigcup_{a_n\in B}\{\omega\colon f(\omega)=a_n\}\\&=\sum_{a_n\in B}A_n\in\mathscr{A}.\end{aligned}$$

最后一个等式求和是对一切使 $a_n\in B$ 的 A_n 求和的. 因此 f 是 $(\Omega,\mathscr{A})$ 上的可测函数. □

现在我们来研究可测函数的构造性质.

定理 5 i) 可测函数是简单函数序列的极限;

ii) 可测函数是初等函数序列的一致极限;

iii) 有界可测函数是简单函数序列的一致极限;

iv) 非负可测函数是非负不降简单函数 (或初等函数) 序列的极限 (相应地, 一致极限).

证 只要对实可测函数证明了定理, 由定理 1 的 3) 可知对复可测函数的情形, 本定理 i) ii) iii) 成立.

设 f 是 $(\Omega,\mathscr{A})$ 上实可测函数.

i) 对任一正整数 n, 任一 $\omega\in\Omega$, 令

$$\begin{aligned}f_n(\omega)=&\sum_{k=-n2^n}^{n2^n-1}\frac{k}{2^n}\chi_{\{\frac{k}{2^n}\leqslant f<\frac{k+1}{2^n}\}}(\omega)+n\chi_{\{f\geqslant n\}}(\omega)\\&-n\chi_{\{f<-n\}}(\omega),\end{aligned}\tag{14}$$

则显然 f_n 是 $(\Omega,\mathscr{A})$ 上的简单函数, 且

$$|f_n(\omega)-f(\omega)|<\frac{1}{2^n}, \text{当} -n\leqslant f(\omega)<n\text{时},$$

$$\begin{aligned}f_n(\omega)&=n, \quad \text{当} f(\omega)=+\infty\text{时},\\ f_n(\omega)&=-n, \quad \text{当} f(\omega)=-\infty\text{时}.\end{aligned}$$

故对一切 $\omega\in\Omega$,

$$\lim_{n\to\infty} f_n(\omega)=f(\omega),$$

因而 i) 获证.

ii) 对任一正整数 n, 任一 $\omega\in\Omega$, 令

$$g_n(\omega)=\sum_{k=-\infty}^{\infty}\frac{k}{2^n}\chi_{\{\frac{k}{2^n}\leqslant f<\frac{k+1}{2^n}\}}(\omega)+\infty\cdot\chi_{\{f=+\infty\}}(\omega)-\infty\cdot\chi_{\{f=-\infty\}}(\omega), \tag{15}$$

则 g_n 是 $(\Omega,\mathscr{A})$ 上初等函数, 且

$$\begin{aligned}|g_n(\omega)-f(\omega)|&<\frac{1}{2^n}, &&\text{当} f(\omega)\neq\pm\infty\text{时},\\ g_n(\omega)&=f(\omega), &&\text{当} f(\omega)=\pm\infty\text{时}.\end{aligned}$$

故对任一正数 $\varepsilon>0$, 有一正整数 N 使 $\frac{1}{2^N}<\varepsilon$, 当 $n>N$ 时, 对一切 $\omega\in\Omega$ 有

$$f(\omega)-\varepsilon\leqslant g_n(\omega)\leqslant f(\omega)+\varepsilon,$$

因此

$$\lim_{n\to\infty} g_n(\omega)=f(\omega), \text{ 对}\omega\in\Omega\text{一致成立}.$$

iii) 当 f 是有界可测函数时, 设 $|f(\omega)|<M$, 对一切 $\omega\in\Omega$ 成立. 则当 $n>M$ 时由 (14) 定义的简单函数 f_n, 最后两项为零, 因而对一切 $\omega\in\Omega$,

$$|f_n(\omega)-f(\omega)|<\frac{1}{2^n}.$$

故

$$\lim_{n\to\infty} f_n(\omega)=f(\omega), \text{对}\omega\in\Omega\text{一致成立}.$$

iv) 当 f 是非负可测函数时, 对一切 $n\geqslant 1$, 任一 $\omega\in\Omega$, 令

$$f_n(\omega)=\sum_{k=0}^{n2^n-1}\frac{k}{2^n}\chi_{\{\frac{k}{2^n}\leqslant f<\frac{k+1}{2^n}\}}(\omega)+n\chi_{\{f\geqslant n\}}(\omega),$$

$$g_n(\omega)=\sum_{k=0}^{+\infty}\frac{k}{2^n}\chi_{\{\frac{k}{2^n}\leqslant f<\frac{k+1}{2^n}\}}(\omega)+\infty\cdot\chi_{\{f=+\infty\}}(\omega).$$

则显然 $\{f_n\}$ 是非降简单函数序列. 事实上, 若 $f_n(\omega)=\frac{k}{2^n}$, 则

$$f_{n+1}(\omega)=\begin{cases}\frac{2k}{2^{n+1}}=\frac{k}{2^n}, \text{当}\omega\in\left\{\frac{2k}{2^{n+1}}\leqslant f<\frac{2k+1}{2^{n+1}}\right\},\\ \frac{2k+1}{2^{n+1}}>\frac{k}{2^n}, \text{当}\omega\in\left\{\frac{2k+1}{2^{n+1}}\leqslant f<\frac{2k+2}{2^{n+1}}\right\},\end{cases}$$

当 $f_n(\omega)=n$ 时, $f_{n+1}(\omega)\geqslant n$, 故 $f_{n+1}\geqslant f_n$, 且 f_n 非负, 仿 i) 知

$$\lim_{n\to\infty}f_n(\omega)=f(\omega), \text{对一切}\omega\in\Omega.$$

即 f 是非负不降简单函数序列的极限.

$\{g_n\}$ 是非负不降初等函数序列, 且仿 ii) 知

$$\lim_{n\to\infty}g_n(\omega)=f(\omega), \text{对一切}\omega\in\Omega\text{一致成立},$$

即 f 是非负不降初等函数序列的一致极限.

定理全部获证. □

定义 5 设 f 是 Ω 上的实函数, 则由

$$f^+(\omega)=\max\{f(\omega),0\}=\begin{cases}f(\omega), & \text{当}\omega\in\{\omega:f(\omega)\geqslant 0\},\\ 0, & \text{当}\omega\in\{\omega:f(\omega)<0\},\end{cases} \tag{16}$$

$$f^-(\omega)=\max\{-f(\omega),0\}=\begin{cases}-f(\omega), & \text{当}\omega\in\{\omega:f(\omega)<0\},\\ 0, & \text{当}\omega\in\{\omega:f(\omega)\geqslant 0\}\end{cases} \tag{17}$$

定义的函数 f^+,f^- 分别称为 f 的**正部**和**负部**. 当 f 是有限值函数时, f^+,f^- 可以表成如下形式:

$$f^+=\frac{|f|+f}{2},\quad f^-=\frac{|f|-f}{2}. \tag{18}$$

易见

$$f=f^+-f^-. \tag{19}$$

定理 6 实可测函数的正部和负部都是可测函数. 因而任何实可测函数可以表成二非负可测函数之差.

证 由 (19) 知只需证明前一部分. 设 f 是 $(\Omega,\mathscr{A})$ 上的实可测函数, f^+ 是 f 的正部, 对任一实数 x,

$$\{\omega:f^+(\omega)<x\}=\begin{cases}\varnothing, & \text{若}x\leqslant 0,\\ \{\omega:f(\omega)<x\}, & \text{若}x>0.\end{cases}$$

故 $\{\omega : f^+(\omega) < x\} \in \mathscr{A}$. 由定理 1 知 f^+ 是 $(\Omega, \mathscr{A})$ 上的可测函数.

$$\{\omega : f^-(\omega) < x\} = \begin{cases} \varnothing, & 若 x \leqslant 0, \\ \{\omega : f(\omega) > -x\}, & 若 x > 0. \end{cases}$$

由于 $(-x, +\infty] \in \tilde{\mathscr{B}}^{(1)}$, 故 $f^{-1}((-x, +\infty]) \in \mathscr{A}$, 即 f^- 是 $(\Omega, \mathscr{A})$ 上的可测函数. □

上面我们讨论了可测函数的构造性质, 我们还将证明任何简单函数序列的极限都是可测函数, 从而给出可测函数的构造性定义. 它在研究测度空间上的积分时起着重要的作用. 为此, 我们研究可测函数对分析运算的封闭性.

2.2.3 可测函数的分析运算

古典分析主要是研究实直线上的连续函数, 或者更一般些, 研究从 $R^{(N)}$ 到 $R^{(N')}$ 的多元连续向量函数. 连续函数在极限运算下已不一定再是连续函数了. 但是我们可以证明 §2.2.1 所定义的可测函数在普通的分析运算下是封闭的. 而且还可以证明, $\tilde{R}^{(n)}$ 或 $\tilde{Z}^{(n)}$ 上的连续函数是 Borel 可测函数.

定义 6 设 $\{a_n\}$ 是一实数序列 (即 $a_n \in \tilde{R}^{(1)}, n \geqslant 1$), 若 $a \in \tilde{R}^{(1)}$ 是 $\{a_n\}$ 的某一子序列 $\{a_{nk}\}$ 的极限, 则称 a 为 $\{a_n\}$ 的一个**极限点**, $\{a_n\}$ 的极限点中最大 (小) 者称为 $\{a_n\}$ 的**上 (下) 极限**, 记作 $\overline{\lim\limits_{n\to\infty}} a_n (\underline{\lim\limits_{n\to\infty}} a_n)$.

设 $\{f_n\}$ 是定义在 Ω 上的实函数序列 (取值在 $\tilde{R}^{(1)}$ 中), 若以 $\sup\limits_n a_n (\inf a_n)$ 表示数序列 $\{a_n\}$ 的上 (下) 确界, 则 $\sup\limits_n f_n (\inf\limits_n f_n)$ 表示对任一给定的 $\omega \in \Omega$ 取值 $\sup\limits_n f_n(\omega)(\inf\limits_n f_n(\omega))$ 的函数, 称为函数序列 $\{f_n\}$ 的**上 (下) 确界**. $\overline{\lim\limits_{n\to\infty}} f_n (\underline{\lim\limits_{n\to\infty}} f_n)$ 表示对任一给定的 $\omega \in \Omega$, 取值 $\overline{\lim\limits_{n\to\infty}} f_n(\omega) (\underline{\lim\limits_{n\to\infty}} f_n(\omega))$ 的函数, 称为函数序列 $\{f_n\}$ 的**上 (下) 极限**. 若 $\overline{\lim\limits_{n\to\infty}} f_n = \underline{\lim\limits_{n\to\infty}} f_n$, 用 $\lim\limits_{n\to\infty} f_n$ 表示, 称为 $\{f_n\}$ 的**极限**.

设 $\{f_n\}$ 是定义在 Ω 上的复函数序列, 且对每一 $n \geqslant 1, f_n = g_n + ih_n, g_n, h_n$ 是 Ω 上的实函数. 若 $\{g_n\}, \{h_n\}$ 的极限都存在, 则称 $\lim\limits_{n\to\infty} g_n + i \lim\limits_{n\to\infty} h_n$ 为 $\{f_n\}$ 的极限, 记作 $\lim\limits_{n\to\infty} f_n$.

由数学分析 (参看 T.M. 菲赫金哥尔茨著的微积分学教程第一卷第一分册 42 段) 可知, 如果上、下极限及上、下确界允许取 $\pm\infty$ 为值, $\underline{\lim\limits_{n\to\infty}} a_n, \overline{\lim\limits_{n\to\infty}} a_n$ 都是存在的, 且

$$\underline{\lim_{n\to\infty}} a_n = \sup_n (\inf_{k\geqslant n} a_k), \overline{\lim_{n\to\infty}} a_n = \inf_n (\sup_{k\geqslant n} a_n),$$

而序列 $\{a_n\}$ 极限存在的充分必要条件是它的上、下极限相等, 且在极限存在时,

$$\lim_{n\to\infty} a_n = \overline{\lim_{n\to\infty}} a_n = \underline{\lim_{n\to\infty}} a_n.$$

由此立得

性质 5 设 $\{f_n\}$ 是定义在 Ω 上的一个实函数序列, 则 $\{f_n\}$ 的上、下确界及上、下极限都存在, 都是 Ω 上的函数, 且

$$\varliminf_{n\to\infty} f_n = \sup_n(\inf_{k\geqslant n} f_k), \varlimsup_{n\to\infty} f_n = \inf_n(\sup_{k\geqslant n} f_k). \tag{20}$$

$\{f_n\}$ 的极限 $\lim\limits_{n\to\infty} f_n$ 存在的充要条件是对一切 $\omega\in\Omega$,

$$\varlimsup_{n\to\infty} f_n(\omega) = \varliminf_{n\to\infty} f_n(\omega) = \lim_{n\to\infty} f_n(\omega). \tag{21}$$

下面我们来证明可测函数对极限运算封闭.

定理 7 i) 设 $\{f_n\}$ 是可测空间 $(\Omega,\mathscr{A})$ 上的一个实可测函数序列, 则 $\sup\limits_n f_n$, $\inf\limits_n f_n$, $\varlimsup\limits_{n\to\infty} f_n$ 以及 $\varliminf\limits_{n\to\infty} f_n$ 都是 $(\Omega,\mathscr{A})$ 上的实可测函数.

ii) 设 $\{f_n\}$ 是 $(\Omega,\mathscr{A})$ 上的复可测函数序列, 且 $\lim\limits_{n\to\infty} f_n(\omega)$, 对一切 $\omega\in\Omega$ 存在, 则 $\lim\limits_{n\to\infty} f_n$ 是 $(\Omega,\mathscr{A})$ 上的可测函数.

证 i) 首先由 $\inf\limits_n a_n$ 的定义知, 若 $\inf\limits_n a_n < x$, 则必有一 $a_n < x$, 反之亦真. 故对任何实数 x

$$\{\omega : \inf_n f_n(\omega) < x\} = \bigcup_n \{\omega : f_n(\omega) < x\} \in \mathscr{A},$$

故 $\inf\limits_n f_n(\omega)$ 是 $(\Omega,\mathscr{A})$ 上的可测函数.

其次, 我们可证: 由 f 可测可知 $-f$ 可测. 因为对任一实数 x,

$$\{\omega : -f(\omega) < x\} = \{\omega : f(\omega) > -x\}.$$

由于 $(-x,+\infty]\in\tilde{\mathscr{B}}^{(1)}$, 且 f 可测, 故 $\{-f<x\}\in\mathscr{A}$. 因而 $-f$ 可测.

第三, 由

$$\sup_n f_n = -\inf_n(-f_n)$$

知 $\sup\limits_n f_n$ 也是可测函数.

第四, 由 (20) 式及上述各结论知 $\varliminf\limits_{n\to\infty} f_n$, $\varlimsup\limits_{n\to\infty} f_n$ 都是可测函数.

总结上述知 i) 获证, 且若 $\lim\limits_{n\to\infty} f_n$ 存在, 则也是可测函数.

ii) 设对一切 $n\geqslant 1, f_n = g_n + ih_n, g_n, h_n$ 是实可测函数, 且 $\lim\limits_{n\to\infty} f_n$ 存在, 则 $\lim\limits_{n\to\infty} g_n$, $\lim\limits_{n\to\infty} h_n$ 存在, 因而由第一部分知它们可测. 故

$$\lim_{n\to\infty} f_n = \lim_{n\to\infty} g_n + i\lim_{n\to\infty} h_n$$

可测. □

推论 可测空间 $(\Omega,\mathscr{A})$ 上的简单函数序列 $\{f_n\}$ 的极限 (如果存在) 仍为可测函数.

由此推论及定理 5 的 i) ii) 可以将 $(\Omega,\mathscr{A})$ 上的可测函数定义为 $(\Omega,\mathscr{A})$ 上收敛的简单函数序列的极限函数, 或 $(\Omega,\mathscr{A})$ 上一致收敛的初等函数序列的极限函数. 这就是可测函数的构造性定义. §2.2.1 中可测函数的定义称为描述性定义. 两种等价定义相辅相成, 各有用场. 在积分理论中构造性定义起着重要的作用; 而为了发现及证明可测函数的一般性质, 描述性定义用起来更加方便.

定理 8　i) 设 g 是定义在 $\tilde{R}^{(n)}$ 的某一子集 D 上的一个函数, 并且它在 D 上沿集合 D 连续. (即对每一 $(x_1,\cdots,x_n)\in D$, 当$(x_1^t,\cdots,x_n^t)\in D$ 且对于 $s=1,\cdots,n,t\to\infty,x_s^t\to x_s$ 时,

$$g(x_1^t,\cdots,x_n^t)\to g(x_1,\cdots,x_n),$$

则称 g**在D上沿集合D连续**.) 则 g 是可测空间 $(D,\tilde{\mathscr{B}}^{(n)}\cap D)$ 上的可测函数.

ii) 在 i) 中用 $\tilde{Z}^{(n)}$ 代替 $\tilde{R}^{(n)}, (z_1,\cdots,z_n)\in D, (z_1^t,\cdots,z_n^t)\in D$ 分别代替 $(x_1,\cdots,x_n)\in D$ 和 $(x_1^t,\cdots,x_n^t)\in D$, 则结论仍然成立.

注　此处所说的连续函数, 比通常数学分析教科书中所指的较广. 因为此处 x_k(或 z_k) 未必是有限值, 而 $g(x_1,\cdots,x_n)$(或 $g(z_1,\cdots,z_n)$) 也未必是有限值.

证　i) 是 ii) 的特殊情形, 按理只需证 ii). 但由于 $\tilde{Z}^{(n)}$ 与 $\tilde{R}^{(2n)}$ 之间存在一一对应且在 $\tilde{Z}^{(n)}$ 的某一子集 D 上的连续函数可以看作相应的 $\tilde{R}^{(2n)}$ 的某一子集 D' 上的连续函数. 而 $(D',\tilde{\mathscr{B}}^{(2n)}\cap D')$ 上的可测函数又可以看作 $(D',\tilde{\mathscr{B}}_z^{(n)}\cap D)$ 上的可测函数. 因而只要证明了 i), ii) 自然成立. 而 i) 的证明在表述上要简单些.

以下往证 i) 成立.

对于每一正整数 k, 把 $\tilde{R}^{(n)}$ 分割成互不相交的边长为 $\dfrac{1}{2^k}$ 的有限维立方体. 即令

$$A_{j_1,\cdots,j_n}=\left[\frac{j_1}{2^k},\frac{j_1+1}{2^k}\right)\times\cdots\times\left[\frac{j_n}{2^k},\frac{j_n+1}{2^k}\right),$$

其中 $j_1,\cdots,j_n$ 为任意整数和 $\pm\infty$. 当某一 j_s 为 $+\infty$ 时, $\left[\dfrac{j_s}{2^k},\dfrac{j_s+1}{2^k}\right)$ 理解为单点集 $\{+\infty\}$, 当某一 j_s 为 $-\infty$ 时, $\left[\dfrac{j_s}{2^k},\dfrac{j_s+1}{2^k}\right)$ 理解为单点集 $\{-\infty\}$. 于是

$$\tilde{R}^{(n)}=\sum_{\substack{j_s=\pm\infty\text{及整数}\\ s=1,\cdots,n}}A_{j_1,\cdots,j_n}.$$

这样的立方体共有可数个, 在每一立方体上每两点的每一坐标相差不到 $\dfrac{1}{2^k}$. 我们把那些与 D 的交不空的立方体进行重排, 记作 $\{A_l^{(k)},l\geqslant 1\}$, 每一 $A_l^{(k)}\in\tilde{\mathscr{B}}^{(n)}$, 因而 $A_l^{(k)}\cap D\in\tilde{\mathscr{B}}^{(n)}\cap D$, 对每一 $l\geqslant 1$, 取定 $(x_{1l}^{(k)},\cdots,x_{nl}^{(k)})\in A_l^{(k)}\cap D$, 对一切

$(x_1,\cdots,x_n)\in D$, 令

$$g_k(x_1,\cdots,x_n)=\sum_{l=1}^{\infty}g(x_{1l}^{(k)},\cdots,x_{nl}^{(k)})\chi_{A_l^{(k)}\cap D}(x_1,\cdots,x_n), \tag{22}$$

则 g_k 是定义在 $(D,\tilde{\mathscr{B}}^{(n)}\cap D)$ 上的初等函数, 因而是可测的.

因为对任一 $(x_1,\cdots,x_n)\in D$, 任一 k, 总存在 l_k, 使 $(x_1,\cdots,x_n)\in A_{l_k}^{(k)}$, 因而对每一 $1\leqslant s\leqslant n$,

$$x_{sl_k}^{(k)}-\frac{1}{2^k}<x_s\leqslant x_{sl_k}^{(k)}+\frac{1}{2^k},\ \text{当 } x_s \text{ 有限时},$$

$$x_s=x_{sl_k}^{(k)},\ \text{当 } x_s \text{ 无限时}.$$

由 (22) 知

$$g_k(x_1,\cdots,x_n)=g(x_{1l_k}^{(k)},\cdots,x_{nl_k}^{(k)}).$$

由于当 $k\to\infty$ 时, 每一 $x_{sl_k}^{(k)}\to x_s, s=1,\cdots,n$. 由题设条件, g 在 D 上沿 D 连续, 因而有

$$\begin{aligned}\lim_{k\to\infty}g_k(x_1,\cdots,x_n)&=\lim_{k\to\infty}g(x_{1l_k}^{(k)},\cdots,x_{nl_k}^{(k)})\\&=\lim_{\substack{x_{sl_k}^{(k)}\to x_s\\ s=1,\cdots,n}}g(x_{1l_k}^{(k)},\cdots,x_{nl_k}^{(k)})=g(x_1,\cdots,x_n)\end{aligned}$$

对一切 $(x_1,\cdots,x_n)\in D$ 成立. 因而 g 作为可测函数序列 $\{g_k\}$ 的极限是可测的. 因而 i) 获证. □

定理 9 设 $f_k,k=1,\cdots,n$ 是 $(\Omega,\mathscr{A})$ 上的复可测函数, 且对每一 $\omega\in\Omega,(f_1(\omega),\cdots,f_n(\omega))\in D\subset\tilde{Z}^{(n)}$. g 是 D 上沿集合 D 的连续函数, 则 $g(f_1,\cdots,f_n)$ 是 $(\Omega,\mathscr{A})$ 上的可测函数.

证 由 $f_1,\cdots,f_n$ 是 $(\Omega,\mathscr{A})$ 上的复可测函数知 $(f_1,\cdots,f_n)$ 是 $(\Omega,\mathscr{A})$ 到 $(\tilde{Z}^{(n)},\tilde{\mathscr{B}}_z^{(n)})$ 的可测映射. 再由对一切 $\omega\in\Omega,(f_1(\omega),\cdots,f_n(\omega))\in D$, 知 $(f_1,\cdots,f_n)$ 是 $(\Omega,\mathscr{A})$ 到 $(D,\tilde{\mathscr{B}}_z^{(n)}\cap D)$ 的可测映射. 再由 g 是 D 上沿集 D 的连续函数及定理 8 知 g 是 $(D,\tilde{\mathscr{B}}_z^{(n)}\cap D)$ 到 $(\tilde{Z}^{(1)},\tilde{\mathscr{B}}_z^{(1)})$ 上的可测映射. 故由定理 2 知 $g(f_1,\cdots,f_n)$ 是 $(\Omega,\mathscr{A})$ 到 $(\tilde{Z}^{(1)},\tilde{\mathscr{B}}_z^{(1)})$ 的可测映射, 即 $g(f_1,\cdots,f_n)$ 是 $(\Omega,\mathscr{A})$ 上的可测函数.

定理 8, 9 有一些重要的特殊情形, 我们把它们写成下列的

推论 1 设 f,g 是可测空间 $(\Omega,\mathscr{A})$ 上的可测函数, 则 $f+g,af$ (a 是任意复数), $f\cdot g$ 及 f/g 都是 $(\Omega,\mathscr{A})$ 上的可测函数.(这里自然应该假定对于每一 $\omega\in\Omega,(f(\omega),g(\omega))$ 分别属于 $h(z_1,z_2)=z_1+z_2,h(z_1,z_2)=z_1\cdot z_2,h(z_1,z_2)=z_1/z_2,f$ 属于 $h(z)=az$ 的连续区域 D 内.)

推论 2 g 是 $(\tilde{R}^{(n)},\tilde{Z}^{(n)})$ 上的连续函数, 则 g 是一 Borel 可测函数.

以上定理及推论说明可测函数对算术运算封闭. Borel 可测函数类包括连续函数类且对极限运算封闭.

由定理 9 及推论 1 我们有

定理 10 设 g 是 $\tilde{Z}^{(n)}$ 上的连续函数, 且在有限点上取有限值, 又设 $\xi_1, \cdots, \xi_n$ 是概率场 $(\Omega, \mathscr{A}, P)$ 上的 n 个随机变量, 则 $g(\xi_1, \cdots, \xi_n)$ 也是 $(\Omega, \mathscr{A}, P)$ 上的随机变量.

特别, 若 ξ, η 是 $(\Omega, \mathscr{A}, P)$ 上的随机变量, 则 $\xi + \eta, a\xi$(a 为任意有限复数), $\xi \cdot \eta$ 都是随机变量.

由这个定理, 我们还可以得出今后将要用到的很多结论. 例如, 若 $(\xi_1, \cdots, \xi_n)$ 是 $(\Omega, \mathscr{A}, P)$ 上的 n 维随机变量, 则 $\dfrac{1}{b}\left(\sum\limits_{k=1}^{n} \xi_k - a\right)$($a, b$为常数, $b \neq 0$), $\dfrac{1}{\sigma^2}\sum\limits_{k=1}^{n}(\xi_k - a_k)^2, e^{i\sum\limits_{k=1}^{n} t_k\xi_k}$ 等都是 $(\Omega, \mathscr{A}, P)$ 上的随机变量.

今后我们还要考虑两个随机变量的商, 为此我们先作一约定: 设 ζ 除去一个零概率集而外有定义且与一个随机变量 ξ 相等, 我们也称 ζ 与 ξ 几乎相等. 在此约定下, 我们有

定理 11 设 (ξ, η) 是概率场 $(\Omega, \mathscr{A}, P)$ 上的随机变量, 且 $P(\eta = 0) = 0$, 则 ξ/η 与一随机变量几乎相等.

证 令

$$\eta_1(\omega) = \begin{cases} \eta(\omega), & 当\eta(\omega) \neq 0, \\ 1, & 当\eta(\omega) = 0. \end{cases}$$

则 η_1 是 $(\Omega, \mathscr{A})$ 上的可测函数. 对任一 $\omega \in \Omega, \eta_1(\omega) \neq 0$, 由定理 8 知 ξ/η_1 是 $(\Omega, \mathscr{A})$ 上的可测函数且取有限值, 故 ξ/η_1 是 $(\Omega, \mathscr{A}, P)$ 上的随机变量. 而 ξ/η 除去一个零概率集 $\{\eta = 0\}$ 而外有定义且与随机变量 ξ/η_1 相等, 于是 ξ/η 与 ξ/η_1 几乎相等. □

2.2.4 函数形式的单调类定理

在 §1.3 我们讨论了集合形式的单调类定理. 这里将要研究的函数形式的单调类定理, 是研究函数性质的一个重要工具. 先给出

定义 7 设 $\mathscr{L}$ 是定义在 Ω 上的一族函数, 满足条件: 若 $f \in \mathscr{L}$, 则 $f^+, f^- \in \mathscr{L}$.

函数族 L 称为 $\mathscr{L}$-系, 如果它满足条件:

I) $1 \in L$(1表示恒等于 1 的函数);

II) L 中有限个函数的线性组合如果有意义则仍属于 L;

III) 如果 $f_n \in L, 0 \leqslant f_n \uparrow f, f$有界或$f \in \mathscr{L}$,则$f \in L$.

定理 12(函数形式的单调类定理). 若 $\mathscr{L}$- 系 L 包含某一 π-系 $\mathscr{C}$ 中任一集 A 的示性函数, 则 L 包含一切属于 $\mathscr{L}$ 的关于 $\sigma(\mathscr{C})$ 可测的实函数.

证 令

$$\Lambda=\{A:\chi_A\in L\}.$$

则由条件 I)— III) 易知, $\Omega\in\Lambda,\Lambda$ 对真差封闭, 对单调增集序列的并封闭, 因而 Λ 是 λ- 系. 由定理条件知, $\Lambda\supset\mathscr{C}$. 故由集合形式的单调类定理知 $\Lambda\supset\sigma(\mathscr{C})$; 换句话说, 对任意集 $A\in\sigma(\mathscr{C}),\chi_A\in L$. 再由 II 知关于 $\sigma(\mathscr{C})$ 可测的简单函数属于 L.

设 f 为 $\mathscr{L}$ 中关于 $\sigma(\mathscr{C})$ 可测的非负函数, 根据定理 5, 存在 $\sigma(\mathscr{C})$ 可测的非负简单函数增序列 $0\leqslant f_n\uparrow f$. 由III) 知 $f\in L$. 若 f 是 $\mathscr{L}$ 中关于 $\sigma(\mathscr{C})$ 可测的, 则 f^+,f^- 都是 $\mathscr{L}$ 中关于 $\sigma(\mathscr{C})$ 可测的, 且 $f^+\in L,f^-\in L,f=f^+-f^-$ 有意义, 根据 II), $f\in L$. 因而定理获证. □

这一定理的典型用法是: 欲证明某一函数族 F 具有某性质 A_0, 为此引入一个函数族 $\mathscr{L}$,($\mathscr{L}$ 满足定义中的条件), 使

$$L=\{f:f\text{具有性质}A_0\}$$

为一 $\mathscr{L}$- 系, 再引进一 π- 系 $\mathscr{C}$, 使 $\mathscr{L}$ 中关于 $\sigma(\mathscr{C})$ 可测的函数类包含 F, 于是根据定理 12 只需证明对一切 $A\in\mathscr{C},\chi_A\in L$ 成立就够了. 这个方法称为 $\mathscr{L}$- 系方法.

下面我们应用 $\mathscr{L}$- 系方法证明几个重要定理.

定理 13 设 Ω 是一集合, $(E,\mathscr{E})$ 是可测空间, f 是 Ω 到 E 的映射. 令 $\sigma(f)=f^{-1}(\mathscr{E})$, 则 φ 是从 Ω 到 $\bar{R}^{(1)}$ 的 $\sigma(f)$ 可测函数的充分必要条件是存在 $(E,\mathscr{E})$ 上实可测函数 g, 使得 $\varphi=g\circ f$, 并且若 φ 有限 (有界), 则可取 g 有限 (有界).

证 如果 $\varphi=g\circ f$, 由于 f 是 $(\Omega,\sigma(f))$ 到 $(E,\mathscr{E})$ 的可测映射, g 是 $(E,\mathscr{E})$ 到 $(\bar{R}^{(1)},\bar{\mathscr{B}}^{(1)})$ 的可测映射, 由定理 2 知 $g\circ f$ 是由 $(\Omega,\sigma(f))$ 到 $(\bar{R}^{(1)},\bar{\mathscr{B}}^{(1)})$ 的可测映射, 即为 $\sigma(f)$ 可测函数条件的充分性获证.

以下往证条件的必要性. 为此, 令 $\mathscr{L}$ 表示 Ω 上一切实函数. 令

$$L=\{g\circ f:g\in\mathscr{E}\}$$

(为叙述简单起见, 今后用 $g\in\mathscr{E}$ 表 g 是 $(E,\mathscr{E})$ 上的可测函数).

首先证明 L 是 $\mathscr{L}$- 系:

$$1=\chi_E\circ f,\chi_E\in\mathscr{E},\text{故}1\in L;$$

L 对线性组合封闭; 事实上, $\alpha(g\circ f)=(\alpha g)\circ f$, 若 $g_1\circ f+g_2\circ f$ 有意义, 令 $A=\{x\in E: g_1(x)=-g_2(x)=\pm\infty\}, g=g_1+g_2\chi_{A^c}\in\mathscr{E}$, 则 $g_1\circ f+g_2\circ f=g\circ f$.

若 $\varphi_n\in L, 0\leqslant\varphi_n\uparrow\varphi$, 则存在 $g_n\in\mathscr{E}$ 使 $\varphi_n=g_n\circ f$, 令 $g=\sup\limits_n g_n$, 则 $g\in\mathscr{E}$ 且 $\varphi=g\circ f$, 因而 $\varphi\in L$.

再证 L 包含 $\sigma(f)$ 中集合的示性函数: 若 $C\in\sigma(f)$, 则存在 $B\in\mathscr{E}$, 使 $C=f^{-1}(B)$, 从而对一切 $\omega\in\Omega, \chi_C(\omega)=\chi_B(f(\omega))$, 即 $\chi_C=\chi_B\circ f$, 而 $\chi_B\in\mathscr{E}$, 故 $\chi_C\in L$. 由定理 12 知 L 包含一切 $\sigma(f)$ 实可测函数.

若 φ 有限, 且 $\varphi=g\circ f$, 令 $g'=g\chi_{\{|g|<\infty\}}$, 则 g' 有限, $g'\in\mathscr{E}$ 且 $\varphi=g'\circ f$. 最后这一等式由于 f 的值域 $\subset\{|g|<\infty\}$.

若 φ 有界, $|\varphi|<M, \varphi=g\circ f$ 令 $g'=g\chi_{\{|g|<M\}}$, 则 g' 有限, $g'\in\mathscr{E}$ 且 $\varphi=g'\circ f$. 最后这一等式由于 $|\varphi|<M$ 知 f 的值域 $\subset\{|g|<M\}$. □

推论 1 $f=(f_1,\cdots,f_n)$ 是 Ω 上的 n 维可测函数, 则 g 关于 $f^{-1}(\tilde{\mathscr{B}}^{(n)})$ 可测的充分必要条件是存在 $(\tilde{R}^{(n)},\tilde{\mathscr{B}}^{(n)})$ 上的可测函数 G, 使对一切 $\omega\in\Omega, g(\omega)=G(f_1(\omega),\cdots,f_n(\omega))$ 成立.

定理 14 设 $\mathscr{L}$ 是定义在 $\tilde{R}^{(n)}$ 上的实函数类, L 是 $\tilde{R}^{(n)}$ 上包含有界连续函数的 $\mathscr{L}-$ 系. 则 L 包含 $\mathscr{L}$ 中一切 Borel 可测函数.

证 令

$$\mathscr{C}=\{A: A\text{是}\tilde{R}^{(n)}\text{中有限开区间或无限区间}\},$$

则 $\mathscr{C}$ 是 $\pi-$ 系, 且 $\sigma(\mathscr{C})=\tilde{\mathscr{B}}^{(n)}$. 故若能证明 $\chi_A\in L$, 对一切 $A\in\mathscr{C}$ 成立, 定理即获证明.

设 $A=\{(x_1,\cdots,x_n), a_k<x_k<b_k, k=1,\cdots,n\}$. 对于任何 $l\geqslant 1, k=1,\cdots,n$, 令

$$f_k^{(1)}(x_k)=\begin{cases}0 & x_k\leqslant a_k,\\ l(x_k-a_k), & a_k<x_k<a_k+\dfrac{1}{l},\\ 1, & a_k+\dfrac{1}{l}\leqslant x_k\leqslant b_k-\dfrac{1}{l},\\ l(b_k-x_k), & b_k-\dfrac{1}{l}<x_k<b_k,\\ 0, & x_k\geqslant b_k,\end{cases}$$

则 $f_k^{(l)}(x_k)$ 看作 $\tilde{R}^{(n)}$ 上的函数是连续、有界且当 $l\to\infty$ 时 $0\leqslant f_k^{(l)}(x_k)\uparrow\chi_{(a_k,b_k)}(\chi_k), k=1,\cdots,n$. 因而 $\prod\limits_{k=1}^{n}f_k^{(l)}(x_k)$ 是非负有界连续函数, 且对一切 $(x_1,\cdots,x_n)\in\tilde{R}^{(n)}$, 有

$$\prod_{k=1}^{n}f_k^{(l)}(x_k)\uparrow\chi_A(x_1,\cdots,x_n).$$

故 $\chi_A \in L$. 若 A 是 $\mathscr{C}$ 中某些分量为无限区间的集, 例如 $A = [-\infty, b_1) \times (a_2, +\infty] \times \cdots \times (a_n, b_n)$, 则用上述方法同样可证 $\chi_A \in L$. 根据定理 12, 本定理获证. □

习题及补充

1. 试证示性函数有下列性质:

i) $\chi_{A\cap B} = \chi_A \chi_B$; $\chi_{A\cup B} = \chi_A + \chi_B - \chi_{A\cap B}$;

$\chi_{A^c} = 1 - \chi_A$; $\chi_{A\Delta B} = |\chi_A - \chi_B|$;

ii) $\chi_{\bigcup\limits_{m=1}^{\infty} A_n} = \sup\limits_n \chi_{A_n}$; $\chi_{\bigcap\limits_{n=1}^{\infty} A_n} = \inf\limits_n \chi_{A_n}$;

$\chi_{\varlimsup\limits_{n\to\infty} A_n} = \varlimsup\limits_{n\to\infty} \chi_{A_n}$; $\chi_{\varliminf\limits_{n\to\infty} A_n} = \varliminf\limits_{n\to\infty} \chi_{A_n}$.

2. 设 f, g 是可测空间 $(\Omega, \mathscr{A})$ 上的实 (或复) 可测函数, 问下列函数是不是 $(\Omega, \mathscr{A})$ 上的可测函数?

i) $f_1(\omega) = \begin{cases} f(\omega), & \omega \in \{|f| < \infty\}, \\ 0, & \omega \in \{|f| = \infty\}; \end{cases}$

ii) $f_2(\omega) = \begin{cases} f(\omega), & \omega \neq \omega_0, \\ f(\omega) + 1, & \omega = \omega_0, \omega_0 \in \Omega; \end{cases}$

iii) $f_3(\omega) = \begin{cases} f(\omega), & \omega \in A, \\ g(\omega), & \omega \notin A, \end{cases} \quad A \in \mathscr{A}.$

3. f 是可测空间 $(\Omega, \mathscr{A})$ 上的实可测函数, 则 $|f|$ 可测, 逆命题是否成立?

4. 若 f, g 是简单函数, $af, f+g, f \cdot g$ 有意义, 试证: $af, f+g, f\cdot g, |f|, f^+$ 及 f^- 都是简单函数.

5. 试求可测空间 $(\Omega, \mathscr{A})$ 上的全部可测函数, 其中 $\mathscr{A} = \{\varnothing, A, A^c, \Omega\}$.

6. 若 f, g 是 $(\Omega, \mathscr{A})$ 上的实可测函数, c 是任意实数, 试证:

$$\{\omega : f(\omega) < g(\omega) + c\} \in \mathscr{A};$$

$$\{\omega : f(\omega) \leqslant g(\omega) + c\} \in \mathscr{A};$$

$$\{\omega : f(\omega) = g(\omega) + c\} \in \mathscr{A}.$$

7. 若 $\{f_n\}$ 是 $(\Omega, \mathscr{A})$ 上实可测函数序列, 则集

$$\{\omega : \varliminf_{n\to\infty} f_n(\omega) = \varlimsup_{n\to\infty} f_n(\omega)\} (\triangleq \{\omega : \lim_{n\to\infty} f_n(\omega)存在\})$$

是可测集.

8. 证明: 若 $f_k, k = 1, \cdots, m$ 是 $R^{(n)}$ 上的 m 个连续函数, 则对任何实数 $y_1, \cdots, y_m$ 来说, 集

$$\{(x_1, \cdots, x_n) : f_k(x_1, \cdots, x_n) < y_k, k = 1, \cdots, m\}$$

是 Borel 集. 并由此说明任何维空间的球 (或椭球)、平面上的多边形、圆柱、圆锥及角锥等图形所围成的区域都是相应空间的 Borel 集.

9. 试证: $g(z)$ 是 $(\tilde{Z}^{(1)}, \tilde{\mathscr{B}}_z^{(1)})$ 上的实可测函数的充分必要条件是 $h(x,y) = g(x+iy)$ 是 $(\tilde{R}^{(2)}, \tilde{\mathscr{B}}^{(2)})$ 上的实可测函数. 将此结论推广到 $g(z_1, \cdots, z_n)$ 是 $(\tilde{Z}^{(n)}, \tilde{\mathscr{B}}_z^{(n)})$ 上的可测函数的情形.

10. 设 $\mathscr{C}$ 为空间 Ω 中的一个 π-系, $\mathscr{H}$ 为 Ω 上的函数类, 如果它满足下述条件:

i) $1 \in \mathscr{H}$;

ii) $\mathscr{H}$ 对非负线性组合封闭, 且若 $f, g \in \mathscr{H}$, 有界, $f \geqslant g$, 则 $f - g \in \mathscr{H}$;

iii) 若 $\{f_n\} \subset \mathscr{H}$, 且 $0 \leqslant f_n \uparrow f$, 则 $f \in \mathscr{H}$;

iv) $\mathscr{H}$ 包含 $\mathscr{C}$ 中集合的示性函数,

则 $\mathscr{H}$ 包含一切关于 $\sigma(\mathscr{C})$ 可测的非负函数.

11. 设 $(\Omega, \bar{\mathscr{A}}, \bar{\mu})$ 是 $(\Omega, \mathscr{A}, \mu)$ 的完全化测度空间, 试证: 对任意 $\bar{\mathscr{A}}-$ 可测函数 f, 必存在一个 $\mathscr{A}-$ 可测函数 f' 使 $f' = f, \text{a.e.}\bar{\mu}$.

12. 设 $f(t,\omega)$ 满足: i) 对每个固定的 $t \in R, f(t,\cdot)$ 是 $(\Omega, \mathscr{A})$ 上的可测函数;ii) 对每个固定的 $\omega \in \Omega, f(\cdot,\omega)$ 是 R 上的连续函数. 试证: f 是乘积空间 $(R\times\Omega, \mathscr{B}^{(1)}\times\mathscr{A})$ 上的可测函数.

$\left(\text{提示: 令} f_n(t,\omega) = \sum\limits_{k=-\infty}^{\infty} f\left(\frac{k+1}{n}, \omega\right)\chi_{[\frac{k}{n}, \frac{k+1}{n})}(t)\right.$, 证明 f_n 是 $\mathscr{B}^{(1)} \times \mathscr{A}$ 可测的, 且 $\left.f_n \to f(n \to \infty)\right)$.

§2.3 分布函数

随机变量是可测空间上取有限值的可测函数. 从概率论的观点主要是研究随机变量取值的概率规律. 我们将会看到, 随机变量的分布函数完全描述了随机变量取值的概率规律. 而且它是我们比较熟悉的欧氏空间上的点函数, 较之抽象空间上的集函数易于掌握.

由于只有实随机变量才有分布函数, 因此将实随机变量一律简称为随机变量. 如无特别声明, 将 $(\xi_1, \cdots, \xi_n)$ 的分布函数一律记作 $F(x_1, \cdots, x_n)$.

今后, 我们不限于研究随机变量的分布函数, 还将研究一般的分布函数, 它与 $(R^{(n)}, \mathscr{B}^{(n)})$ 上对有界集取有限值的测度相对应.

在本节中使用了多元函数的差分运算符号. 对于函数 $F(x_1, \cdots, x_n)$, 令

$$\begin{aligned}\Delta_{b_k,a_k}^{x_k} F = & F(x_1, \cdots, x_{k-1}, b_k, x_{k+1}, \cdots, x_n) \\ & -F(x_1, \cdots, x_{k-1}, a_k, x_{k+1}, \cdots, x_n),\end{aligned} \tag{1}$$

$$\Delta_{b_{k_l},a_{k_l}}^{x_{k_l}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F=\Delta_{b_{k_l},a_{k_l}}^{x_{k_l}}[\Delta_{b_{k_{l-1}},a_{k_{l-1}}}^{x_{k_{l-1}}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F], \tag{2}$$

其中 $k_1,\cdots,k_l$ 是 $1,\cdots,n$ 中 l 个不同的数;

$$\Delta_{b,a}F=\Delta_{b_n,a_n}^{x_n}\cdots\Delta_{b_1,a_1}^{x_1}F, \text{其中} b=(b_1,\cdots,b_n),$$

$$a=(a_1,\cdots,a_n). \tag{3}$$

由 (1)—(3) 知 F 经过 l 阶差分运算后, 可以用 F 在某些点的值的代数和具体表出. 且在 (2) 中将 $k_1,\cdots,k_l$ 的次序任意调换后所得的结果是一样的. 而且还有

$$\Delta_{b_{k_l},a_{k_l}}^{x_{k_l}}\cdots\Delta_{b_{k_i},c_{k_i}}^{x_{k_i}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F+\Delta_{b_{k_l},a_{k_l}}^{x_{k_l}}\cdots\Delta_{c_{k_i},a_{k_i}}^{x_{k_i}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F$$

$$=\Delta_{b_{k_l},a_{k_l}}^{x_{k_l}}\cdots\Delta_{b_{k_i},a_{k_i}}^{x_{k_i}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F. \tag{4}$$

2.3.1 随机变量的分布函数的性质

性质 1 $F(x_1,\cdots,x_n)$ 对每个自变量 x_k 是不降的, 即对一切 $1\leqslant k\leqslant n, b_k\geqslant a_k$, 有

$$\Delta_{b_k,a_k}^{x_k}F\geqslant 0. \tag{5}$$

证 由分布函数的定义及概率的非负性可知

$$\begin{aligned}\Delta_{b_k,a_k}^{x_k}F&=F(\cdots x_{k-1},b_k,x_{k+1},\cdots)\\&\quad -F(\cdots,x_{k-1},a_k,x_{k+1},\cdots)\\&=P(\cdots,\xi_{k-1}<x_{k-1},\xi_k<b_k,\xi_{k+1}<x_{k+1},\cdots)\\&\quad -P(\cdots,\xi_{k-1}<x_{k-1},\xi_k<a_k,\xi_{k+1}<x_{k+1},\cdots)\\&=P(\cdots,\xi_{k-1}<x_{k-1},a_k\leqslant\xi_k<b_k,\xi_{k+1}<x_{k+1},\cdots)\geqslant 0.\end{aligned}$$

□

当存在 k, 使 $a_k=b_k$ 时, $\Delta_{ba}F=0$.

性质 2 设 $a_{k_m}\leqslant b_{k_m}, m=1,\cdots,l, k_1,\cdots,k_l$ 是 $1,\cdots,n$ 中 l 个不同的数, 则

$$P\left(\bigcap_{m=1}^{l}\{a_{k_m}\leqslant\xi_{k_m}<b_{k_m}\}\bigcap_{h\neq k_m}\{\xi_h<x_h\}\right)$$

$$=\Delta_{b_{k_l},a_{k_l}}^{x_{k_l}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F,\quad l=1,\cdots,n. \tag{6}$$

因而

$$\Delta_{b_{k_l},a_{k_l}}^{x_{k_l}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F\geqslant 0. \tag{7}$$

特别若 $a_k\leqslant b_k, k=1,\cdots,n$, 则 $\Delta_{b,a}F\geqslant 0$. (8)

证 只要证明 (6) 式成立, 其他结论立刻得到. 由性质 1 知当 $l=1$ 时 (6) 式成立. 设 (6) 对 $l-1$ 成立, 则由概率的性质、$\{a_{k_l}\leqslant\xi_{k_l}<b_{k_l}\}=\{\xi_{k_l}<b_{k_l}\}-\{\xi_{k_l}<a_{k_l}\}$ 及 (6) 对 $l-1$ 的情形成立知 (6) 对 l 的情形的左端为

$$
\begin{aligned}
&P\left(\bigcap_{m=1}^{l-1}\{a_{k_m}\leqslant\xi_{k_m}<b_{k_m}\}\cap\{\xi_{k_l}<b_{kl}\}\bigcap_{h\neq k_m}\{\xi_h<x_h\}\right)\\
&-P\left(\bigcap_{m=1}^{l-1}\{a_{k_m}\leqslant\xi_{k_m}<b_{k_m}\}\cap\{\xi_{k_l}<a_{k_l}\}\bigcap_{h\neq k_m}\{\xi_h<x_h\}\right)\\
=&\left[\Delta_{b_{k_{l-1}},a_{k_{l-1}}}^{x_{k_{l-1}}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F\right]_{x_{k_l}=b_{k_l}}\\
&-\left[\Delta_{b_{k_{l-1}},a_{k_{l-1}}}^{x_{k_{l-1}}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F\right]_{x_{k_l}=a_{k_l}}\\
=&\Delta_{b_{k_l},a_{k_l}}^{x_{k_l}}\left[\Delta_{b_{k_{l-1}},a_{k_{l-1}}}^{x_{k_{l-1}}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F\right]\\
=&\Delta_{b_{k_l},a_{k_l}}^{x_{k_l}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F.
\end{aligned}
$$

此即 (6) 式对 l 的情形也成立. 其中 $[\quad]_{x_{k_l}=x}$ 表示括号中的 x_{k_l} 以 x 代替后所得到的表示式. □

性质 3 设 $k_1,\cdots,k_l$ 是 $1,\cdots,n$ 中 l 个不同的数, $b_{k_1},\cdots,b_{k_l}$ 是任意给定的 l 个实数, 则

$$\lim_{\substack{a_{k_i}\to b_{k_i}-0\\ i=1,\cdots,l}}\Delta_{b_{k_l},a_{k_l}}^{x_{k_l}}\cdots\Delta_{b_{k_1},a_{k_1}}^{x_{k_1}}F=0 \tag{9}$$

对一切实数 $x_h(h\neq k_l,\cdots,k_l)$ 成立, 因而 $F(x_1,\cdots,x_n)$ 对每一自变量是左连续的, 即对任一 $k(1\leqslant k\leqslant n)$

$$\lim_{x'_k\to x_k-0}F(\cdots x_{k-1},x'_k,x_{k+1},\cdots)=F(\cdots x_{k-1},x_k,x_{k+1},\cdots). \tag{10}$$

证 令 $\{\varepsilon_m\}$ 为趋于零的递减正数序列, 我们先证

$$\lim_{m\to\infty}\Delta_{b_{k_l},b_{k_l}-\varepsilon_m}^{x_{k_l}}\cdots\Delta_{b_{k_1},b_{k_1}-\varepsilon_m}^{x_{k_1}}F=0. \tag{11}$$

由 (6) 知

$$
\begin{aligned}
&\Delta_{b_{k_l},b_{k_l}-\varepsilon_m}^{x_{k_l}}\cdots\Delta_{b_{k_1},b_{k_1}-\varepsilon_m}^{x_{k_1}}F\\
=&P\left(\bigcap_{i=1}^{l}\{b_{k_i}-\varepsilon_m\leqslant\xi_{k_i}<b_{k_i}\}\bigcap_{h\neq k_i}\{\xi_h<x_h\}\right).
\end{aligned} \tag{12}
$$

记上式右端括号中的事件为 A_m, 则由 $\varepsilon_m \downarrow$ 知$A_m \downarrow \bigcap\limits_{m=1}^{\infty} A_m$, 再由 $\varepsilon_m \to 0$ 易证 $\bigcap\limits_{m=1}^{\infty} A_m = \varnothing$, 因此由概率的连续性及 (12) 式知 (11) 式的左端等于

$$\lim_{m\to\infty} P(A_m) = P\left(\bigcap_{m=1}^{\infty} A_m\right) = 0.$$

今再应用 (11) 式来证明 (9) 式. 设 ε 为任意正数, 则由 (11) 式知必有一 ε_m 存在, 使

$$0 \leqslant \Delta^{x_{k_l}}_{b_{k_l}, b_{k_l}-\varepsilon_m} \cdots \Delta^{x_{k_1}}_{b_{k_1}, b_{k_1}-\varepsilon_m} F < \varepsilon. \tag{13}$$

于是当 $b_{k_i} - \varepsilon_m < a_{k_i} < b_{k_i}, 1 \leqslant i \leqslant l$ 时由 (6), (12) 及 (13) 知

$$0 \leqslant \Delta^{x_{k_l}}_{b_{k_l}, a_{k_l}} \cdots \Delta^{x_{k_1}}_{b_{k_1}, a_{k_1}} F$$

$$= P\left(\bigcap_{i=1}^{l} \{a_{k_i} \leqslant \xi_{k_i} < b_{k_i}\} \bigcap_{h\neq k_i} \{\xi_h < x_h\}\right) \leqslant P(A_m) < \varepsilon,$$

其中用到事件的关系式

$$\bigcap_{i=1}^{l} \{a_{k_i} \leqslant \xi_{k_i} < b_{k_i}\} \bigcap_{h\neq k_i} \{\xi_h < x_h\} \subset A_m$$

及概率的不降性. 故 (9) 式获证.

在 (9) 式中, 令 $l = 1, k_1 = k, a_k = x'_k, b_k = x_k$, 即得 (10) 式. □

性质 4 $F(x_1, \cdots, x_n)$ 还满足

$$F(\cdots, x_{k_1-1}, -\infty, x_{k_1+1}, \cdots x_{k_{l-1}}, -\infty, x_{k_{l+1}}, \cdots) = 0, \tag{14}$$

$$F_{\xi_{k_1}, \cdots \xi_{k_l}}(x_{k_1}, \cdots, x_{k_l})$$
$$= F(\cdots, \infty, x_{k_1}, \infty, \cdots \infty, x_{k_l}, \infty, \cdots), \tag{15}$$

$$F(\infty, \cdots, \infty) = 1, \tag{16}$$

其中 $k_1, \cdots, k_l$ 是 $1, \cdots, n$ 中的 l 个不同的数, 而 (14) 的左端、(15) 的右端及 (16) 的左端分别表示下列三极限:

$$\lim_{\substack{x_{k_m}\to-\infty \\ m=1,\cdots,l}} F(x_1, \cdots, x_n), \quad \lim_{\substack{x_h\to\infty \\ h\neq k_m \\ m=1,\cdots,l}} F(x_1, \cdots, x_n) \text{及}$$

$$\lim_{\substack{x_k \to \infty \\ k=1,\cdots,n}} F(x_1, \cdots, x_n).$$

证 (14)—(16) 的证明与性质 3 的证明完全类似. 我们只证 (15), 其他请读者自证.

为了证明 (15), 我们先证: 若 $\{r\}$ 是正整数序列, 则

$$\begin{aligned} &F_{\xi_{k_1}\cdots\xi_{k_l}}(x_{k_1}, \cdots, x_{k_l}) \\ &= \lim_{r \to \infty} F(\cdots r, x_{k_1}, r, \cdots, r, x_{k_l}, r, \cdots). \end{aligned} \tag{17}$$

为此令

$$B_r = \bigcap_{m=1}^{l} \{\xi_{k_m} < x_{k_m}\} \bigcap_{k \neq k_m} \{\xi_k < r\},$$

则易见 $B_1 \subset B_2 \subset \cdots$, 且由 ξ_k 的值有限知

$$\bigcup_{r=1}^{\infty} B_r = \bigcap_{m=1}^{l} \{\xi_{k_m} < x_{k_m}\}$$

于是由分布函数的定义及概率的连续性知 (17) 的右端等于

$$\lim_{r \to \infty} P(B_r) = P\left(\bigcap_{m=1}^{l} \{\xi_{k_m} < x_{k_m}\}\right) = F_{\xi_{k_1},\cdots\xi_{k_l}}(x_{k_1}, \cdots, x_{k_l}).$$

因而 (17) 式获证.

设 ε 是任一正数, 则由 (17) 知必有 $-r$ 使

$$P(B_r) \leqslant F_{\xi_{k_1},\cdots,\xi_{k_l}}(x_{k_1}, \cdots, x_{k_l}) < P(B_r) + \varepsilon. \tag{18}$$

则当 $x_k > r, k \neq k_1, \cdots, k_l$ 时, $B_r \subset \bigcap\limits_{k=1}^{n} \{\xi_k < x_k\} \subset \bigcap\limits_{m=1}^{l} \{\xi_{k_m} < x_{k_m}\}$, 故由 (18) 知

$$P(B_r) \leqslant P\left(\bigcap_{k=1}^{n} \{\xi_k < x_k\}\right) = F(x_1, \cdots, x_n) \leqslant F_{\xi_{k_1}\cdots\xi_{k_l}}(x_{k_1}, \cdots, x_{k_l}) < P(B_r) + \varepsilon, \tag{19}$$

于是由 (19) 知

$$0 \leqslant F_{\xi_{k_1},\cdots,\xi_{k_l}}(x_{k_1}, \cdots, x_{k_l}) - F(x_1, \cdots, x_n) < \varepsilon,$$

因而 (15) 式获证. □

性质 4 中的 (15) 式说明, 如果我们知道 $(\xi_1, \cdots, \xi_n)$ 的分布函数, 那么可以应用 (15) 求出它的一部分随机变量 $\xi_{k_1}, \cdots, \xi_{k_l}$ 作成的 l 维随机变量的分布函数. 今举一例来说明.

例 1 设 (ξ_1,ξ_2) 是按二维正态 $N(a_1,a_2,\sigma_1^2,\sigma_2^2,r)$ 分布的随机变量, 且 $|r|<1$, 则 ξ_1,ξ_2 分别按 $N(a_1,\sigma_1^2),N(a_2,\sigma_2^2)$ 分布.

证 周知 (ξ_1,ξ_2) 的分布函数是

$$F(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\exp\left\{-\frac{1}{2(1-r^2)}\right.$$
$$\left.\times\left[\frac{(u-a_1)^2}{\sigma_1^2}-2r\frac{(u-a_1)(v-a_2)}{\sigma_1\sigma_2}+\frac{(v-a_2)^2}{\sigma_2^2}\right]\right\}dudv.$$

而

$$\begin{aligned}&-\frac{1}{2(1-r^2)}\left[\frac{(u-a_1)^2}{\sigma_1^2}-2r\frac{(u-a_1)(v-a_2)}{\sigma_1\sigma_2}+\frac{(v-a_2)^2}{\sigma_2^2}\right]\\=&-\frac{(u-a_1)^2}{2\sigma_1^2}-\frac{1}{2(1-r^2)}\left[\frac{v-a_2}{\sigma_2}-r\frac{(u-a_1)}{\sigma_1}\right]^2\\=&-\frac{(u-a_1)^2}{2\sigma_1^2}-\frac{1}{2(1-r^2)\sigma_2^2}\left[v-a_2-\frac{r\sigma_2(u-a_1)}{\sigma_1}\right]^2,\end{aligned}$$

故由 (15) 知

$$\begin{aligned}F_{\xi_1}(x_1)&=\lim_{x_2\to\infty}F(x_1,x_2)\\&=\frac{1}{\sqrt{2\pi}\sigma_1}\int_{-\infty}^{x_1}e^{-\frac{(u-a_1)^2}{2\sigma_1^2}}\times\left[\frac{1}{\sigma_2\sqrt{2\pi(1-r^2)}}\times\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2(1-r^2)\sigma_2^2}\right.\right.\\&\quad\left.\left.\times\left[v-a_2-\frac{r\sigma_2(u-a_1)}{\sigma_1}\right]^2\right\}dv\right]du.\end{aligned}$$

对于任一给定的 u 来说, 令

$$t=\frac{1}{\sqrt{2(1-r^2)}\sigma_2}\left[v-a_2-\frac{r\sigma_2(u-a_1)}{\sigma_1}\right],$$

则易知上式方括号中的积分等于

$$\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}e^{-t^2}dt=\frac{2}{\sqrt{\pi}}\int_0^{\infty}e^{-t^2}dt=1,$$

从而可知 ξ_1 是按 $N(a_1,\sigma_1^2)$ 分布的. 同样可证 ξ_2 是按 $N(a_2,\sigma_2^2)$ 分布的.

总结以上性质, 可得

定理 1 设 $F(x_1,\cdots,x_n)$ 是 n 维随机变量 $(\xi_1,\cdots,\xi_n)$ 的分布函数, 则 F 具有下列性质:

(a) $\Delta_{b,a}F\geqslant 0$, 其中 $a=(a_1,\cdots,a_n),b=(b_1,\cdots,b_n),a_k\leqslant b_k,k=1,\cdots,n$;

(b) $F(x_1,\cdots,x_n)$ 对每一自变量左连续;

(c) 对于任意 $1\leqslant k_1<\cdots<k_l\leqslant n(l\geqslant 1), F(\cdots,x_{k_1-1},-\infty,x_{k_1+1},\cdots,x_{k_l-1},$ $-\infty,x_{k_l+1},\cdots)=0$;

(d) $F(\infty,\cdots,\infty)=1$.

我们将要证明任一满足条件 (a), (b), (c), (d) 的 $R^{(n)}$ 上的函数必为某一 n 维随机变量的分布函数. (即定理 1 的逆定理.) 这个结论的证明是通过以下两个步骤完成的: 第一步证明满足 (a)—(d) 的 n 元函数决定一个概率场 $(R^{(n)},\mathscr{B}^{(n)},P)$; 第二步证明在这个概率场上定义 n 个随机变量 $\xi_k(x_1,\cdots,x_n)=x_k, k=1,\cdots,n$, 则 $(\xi_1,\cdots,\xi_n)$ 的分布函数就是给定的函数. 关于第一步, 事实上可以证明更广一点的定理, 即证明满足 (a), (b) 的任意有限函数可以唯一地决定一个测度空间 $(R^{(n)},\mathscr{B}^{(n)},\mu)$. 为了以后的应用, 我们宁愿证明这个广一点的结论. 证明的方法与 §1.5 建立直线上 Lebesgue 测度的方法在实质上是一样的.

2.3.2 分布函数与 L–S 测度

定义 1 设 $F(x_1,\cdots,x_n)$ 是定义在 $R^{(n)}$ 上的有限值函数, 且满足

(a) $\Delta_{b,a}F\geqslant 0$, 其中 $a=(a_1,\cdots,a_n), b=(b_1,\cdots,b_n), a,b\in R^{(n)}$, 且对一切 $k=1,\cdots,n, a_k\leqslant b_k$;

(b) $F(x_1,\cdots,x_n)$ 对每一自变量左连续, 则称 F 是 $R^{(n)}$ 上的分布函数.

现在我们来叙述并证明本节的主要定理, 即 $R^{(n)}$ 上的分布函数决定 $(R^{(n)},\mathscr{B}^{(n)})$ 上测度的.

定理 2 设 F 是定义在 $R^{(n)}$ 上的一个分布函数, 则在 $\mathscr{B}^{(n)}$ 上存在唯一的测度 μ_F 满足

$$\mu_F([a,b))=\Delta_{b,a}F, \text{对一切} a,b\in R^{(n)}, a\leqslant b, \tag{20}$$

μ_F 在有限区间上取有限值, 一般称为 Borel-Stieltjes 测度, 同时也可以唯一决定 μ_F^*- 可测集类上的测度, 称它为 Lebesgue-Stieltjes 测度 (简称 L–S 测度). 在一般情况下, 两者统称为 L–S 测度.

我们通过以下几个引理来证明定理 2.

引理 1 设

$$\mathfrak{E}^{(n)}=\left\{\begin{array}{c} a=(a_1,\cdots,a_n), a_k\in R \text{ 或 } a_k=-\infty, \\ [\mathrm{a},\mathrm{b}), \text{ 当 } a_k=-\infty \text{ 时}, [a_k,b_k)\triangleq(-\infty,b_k), \\ b=(b_1,\cdots,b_n), b_k\in R \text{ 或 } b_k=+\infty. \end{array}\right\}$$

则 $\mathfrak{E}^{(n)}$ 是 $R^{(n)}$ 中的半集代数.

证 i) $\varnothing=[a,a)\in\mathfrak{E}^{(n)}, R^{(n)}=(-\infty^{(n)},\infty^{(n)})\in\mathfrak{E}^{(n)}$ 其中 $a=(a_1,\cdots,a_n)$, $-\infty^{(n)}=(-\infty,\cdots,-\infty), \infty^{(n)}=(\infty,\cdots,\infty)$.

ii) 若 $A,B\in\mathfrak{C}^{(n)}, A=[a^{(1)},b^{(1)}), B=[a^{(2)},b^{(2)}), a^{(i)}=(a_1^{(i)},\cdots,a_n^{(i)}), b^{(i)}=(b_1^{(i)},\cdots,b_n^{(i)}), i=1,2$. 则令 $c_k=\max\{a_k^{(1)},a_k^{(2)}\}, d_k=\min\{b_k^{(1)},b_k^{(2)}\}, k=1,\cdots,n$, $c=(c_1,\cdots,c_n), d=(d_1,\cdots,d_n)$, 于是 $A\cap B=[c,d)\in\mathfrak{C}^{(n)}$.

iii) 若 $A\in\mathfrak{C}^{(n)}, A=[a,b), a=(a_1,\cdots,a_n), b=(b_1,\cdots,b_n)$, 设 $A_k^{(1)}=(-\infty,a_k), A_k^{(2)}=[a_k,b_k), A_k^{(3)}=[b_k,\infty)$, 则 $A_1^{(i_1)}\times\cdots\times A_n^{(i_n)}\in\mathfrak{C}^{(n)}$, 其中 $i_k=1,2,3, k=1,\cdots,n$ 两两不交, $A_1^{(i_1)}\times\cdots\times A_n^{(i_n)}=\{(x_1,\cdots,x_n): x_k\in A_k^{(i_k)}, k=1,\cdots,n\}, R^{(n)}=\sum\limits_{i_1=1}^{3}\cdots\sum\limits_{i_n=1}^{3}A_1^{(i_1)}\times\cdots\times A_n^{(i_n)}$. 故 A^c 可表示为 $\mathfrak{C}^{(n)}$ 中两两不交的有限个集合的并 (至多 3^n-1 个).

由以上三点可知 $\mathfrak{C}^{(n)}$ 是 $R^{(n)}$ 中的半集代数. □

引理 2 设 $F(x_1,\cdots,x_n)$ 满足定理 2 的条件, 若 $a,b\in R^{(n)}, a\leqslant b, A=[a,b)$, 令

$$\mu(A)\triangleq\mu([a,b))=\Delta_{b,a}F. \tag{21}$$

若 $A=[a,b)$ 是 $R^{(n)}$ 中的无限区间时, 令 $I^{(N)}=[-N,N)=\{(x_1,\cdots,x_n): -N\leqslant x_k<N, k=1,\cdots,n\}$. 令

$$\mu(A)\triangleq\lim_{n\to\infty}\mu([a,b)\cap I^{(N)}). \tag{22}$$

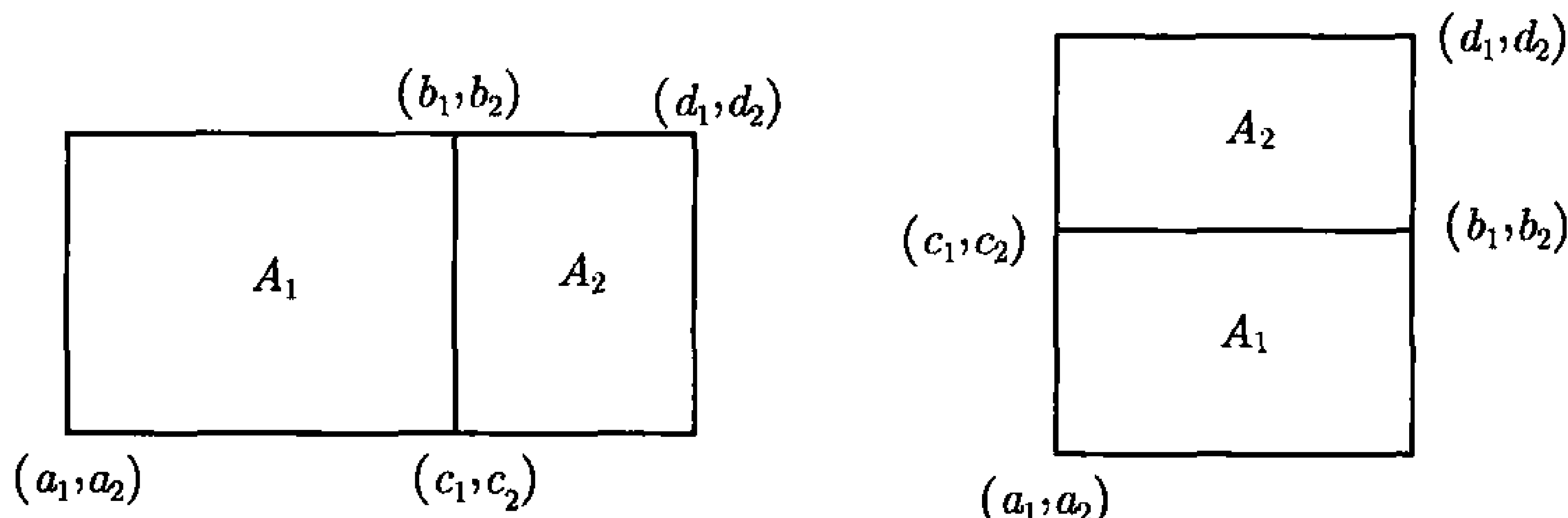
则 μ 是定义在 $\mathfrak{C}^{(n)}$ 上的有限可加测度.

证 i) 先证明在一切有限左闭右开区间类上由 (21) 定义的集函数是有限可加的. 即设 $A_k\in\mathfrak{C}^{(n)}, k=1,\cdots,l$ (l 是任意正整数), 是两两不交的有限区间, 且 $A=\sum\limits_{k=1}^{l}A_k\in\mathfrak{C}^{(n)}$, 则

$$\mu(A)=\sum_{k=1}^{l}\mu(A_k). \tag{23}$$

若 $A=\varnothing$,(23) 式显然成立, 对 $A\neq\varnothing$, 的情形往证 (23) 式成立.

应用数学归纳法来证明. 在证明之前, 先就二维的情形解释一下证明的想法. 首先设 $l=2$, 此时 A_1,A_2 的图形的位置必为



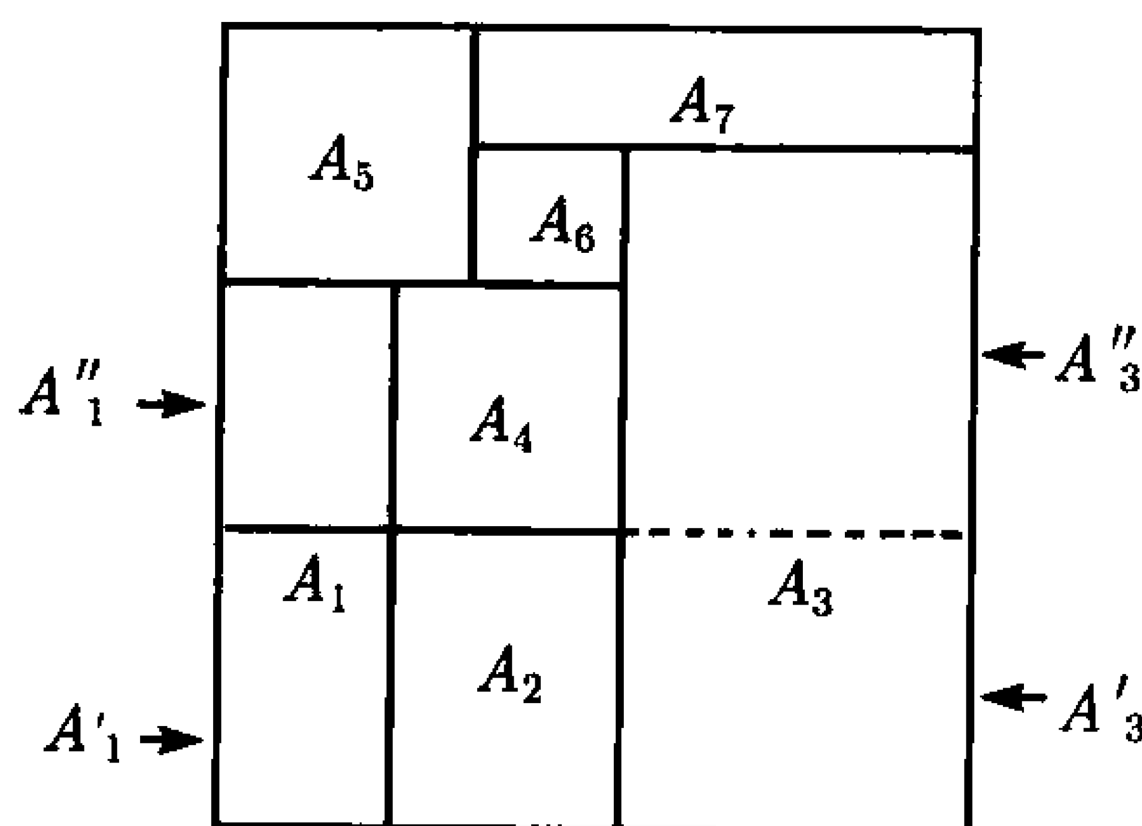
上面二图之一所示. 即若 $A_1=[a,b), A_2=[c,d)$, 则$b_1=c_1, b_2=d_2$且$a_2=c_2$(或$b_2=c_2, a_1=c_1$且$b_1=d_1$), 于是由 (4) 式知

$$\begin{aligned}\mu(A_1)+\mu(A_2) &= \Delta_{b_1,a_1}^{x_1}\Delta_{b_2,a_2}^{x_2}F+\Delta_{d_1,c_1}^{x_1}\Delta_{d_2,c_2}^{x_2}F \\ &= \Delta_{c_1,a_1}^{x_1}\Delta_{d_2,a_2}^{x_2}F+\Delta_{d_1,c_1}^{x_1}\Delta_{d_2,a_2}^{x_2}F \\ &= \Delta_{d_1,a_1}^{x_1}\Delta_{d_2,a_2}^{x_2}F=\Delta_{d,a}F \\ &= \mu(A_1+A_2).\end{aligned}$$

(上右图的情形同样证明).$l=2$ 的情形获证.

对于一般的情形 $A_1,\cdots,A_l$ 的图形的位置如下图所示 (图上画的是 $l=7$ 的情形). 如图画一虚线, 将 A_1 分成 A_1', A_1'', 将 A_3 分成 A_3', A_3''. 于是由 $i=2$ 的情形知 $\mu(A_1)=\mu(A_1')+\mu(A_1''), \mu(A_3)=\mu(A_3')+\mu(A_3'')$. 设 (23) 对 $l<7$ 的情形成立, 则由 (23) 对 $l=3, l=6$ 及 $l=2$ 的情形如

$$\begin{aligned}\sum_{k=1}^{7}\mu(A_k) &= [\mu(A_1')+\mu(A_2)+\mu(A_3')]+\left[\mu(A_1'')+\mu(A_3'')+\sum_{k=4}^{7}\mu(A_k)\right] \\ &= \mu(A_1'+A_2+A_3')+\mu(A_1''+A_3''+\sum_{k=4}^{7}A_k) \\ &= \mu\left(A_1'+A_2+A_3'+A_1''+A_3''+\sum_{k=4}^{7}A_k\right) \\ &= \mu\left(\sum_{k=4}^{7}A_k\right).\end{aligned}$$



下面给出 (23) 式的证明.

当 $l=2$ 时, 我们先来考虑 A_1, A_2 的图形性质, 设 $A_1=[a,b)\neq\varnothing, A_2=[c,d)\neq\varnothing, A=[e,f)$, 则 $A=A_1+A_2$ 可写成

$$\{(x_1,\cdots,x_n): a_k\leqslant x_k<b_k, k=1,\cdots,n\}$$

$$+\{(x_1,\cdots,x_n):c_k\leqslant x_k<d_k,k=1,\cdots,n\}$$
$$=\{(x_1,\cdots,x_n):e_k\leqslant x_k<f_k,k=1,\cdots,n\}. \tag{24}$$

于是易知

$$[a_k,b_k)\cup[c_k,d_k)=[e_k,f_k),k=1,\cdots,n. \tag{25}$$

由于 $A_1\cap A_2=\varnothing$, 故必有一 h 使 $[a_h,b_h)$ 与 $[c_h,d_h)$ 不相交, 由 (25) 知必有 $b_h=c_h$ 或 $a_h=d_h$. 不妨设

$$b_h=c_h,\text{因而}a_h=e_h,d_h=f_h. \tag{26}$$

今再证

$$[a_k,b_k)=[c_k,d_k)=[e_k,f_k),k\neq h.k=1,\cdots,n. \tag{27}$$

假设 (27) 不成立, 则必有一 $i\neq h$, 使 $[a_i,b_i)\neq[c_i,d_i)$, 于是有一 $\bar{x}_i\in[e_i,f_i)$ 使 $\bar{x}_i\in[a_i,b_i)-[c_i,d_i)$(或 $\bar{x}_i\in[c_i,d_i)-[a_i,b_i)$), 令

$$x=(x_1,\cdots,x_n)\in[e,f),x_h=c_h,x_i=\bar{x}_i.$$

(或相应地 $x_h=a_h,x_i=\bar{x}_i$.) 则 $x\in A$, 但 $x\notin A_1+A_2$, 因为由 (25), (26) 知 $x_h\notin[a_h,b_h)$(或 $x_h\notin[a_h,d_h)$), $x_i\notin[c_i,d_i)$(或 $x_i\notin[a_i,b_i)$), 故 $x\notin A_1+A_2$, 这与 $A=A_1+A_2$ 矛盾, 因而. (27) 式成立. 于是由 (21), (26), (27) 及 (4) 式知

$$\begin{aligned}\mu(A_1)+\mu(A_2)&=\Delta_{b,a}F+\Delta_{d,c}F\\&=\Delta_{b_1,a_1}^{x_1}\cdots\Delta_{b_h,a_h}^{x_h}\cdots\Delta_{b_n,a_n}^{x_n}F+\Delta_{d_1,c_1}^{x_1}\cdots\Delta_{d_h,c_h}^{x_h}\cdots\Delta_{d_n,c_n}^{x_n}F\\&=\Delta_{f_1,e_1}^{x_1}\cdots\Delta_{c_h,e_h}\cdots\Delta_{f_n,e_n}F+\Delta_{f_1,e_1}^{x_1}\cdots\Delta_{f_h,c_h}^{x_h}\cdots\Delta_{f_n,e_n}^{x_n}F\\&=\Delta_{f_1,e_1}^{x_1}\cdots\Delta_{f_h,e_h}^{x_h}\cdots\Delta_{f_n,e_n}^{x_n}=\Delta_{f,e}F=\mu(A).\end{aligned}$$

故 (23) 式对 $l=2$ 的情形成立.

设 (23) 对 $l<m(m>1)$ 的情形成立, 今证 (23) 对 $l=m$ 成立. 设 $A\in\mathfrak{E}^{(n)}$; $A_r\in\mathfrak{E}^{(n)},r=1,\cdots,m$, 为两两不交的有限区间, 且 $A=\sum\limits_{r=1}^{m}A_r$, 由 $A_1\cap A_2=\varnothing$, 设

$$A_i=[a^{(i)},b^{(i)}),a^{(i)}=(a_1^{(i)},\cdots,a_n^{(i)}),$$
$$b^{(i)}=(b_1^{(i)},\cdots,b_n^{(i)}),i=1,2.$$

必存在一个 $k(1\leqslant k\leqslant n)$ 使得 $[a_k^{(1)},b_k^{(1)})\cap[a_k^{(2)},b_k^{(2)})=\varnothing$, 不妨设

$$a_k^{(1)}\leqslant b_k^{(1)}\leqslant a_k^{(2)}\leqslant b_k^{(2)},$$

取

$$B=\{(x_1,\cdots,x_n):x_k<b_k^{(1)},x_h\in R,h\neq k\}\in\mathfrak{E}^{(n)}.$$

则 $B^c=\{(x_1,\cdots,x_n):x_k\geqslant b_k^{(1)},x_h\in R,h\neq k\}\in\mathfrak{E}^{(n)}$. 于是 $B\cap A,B^c\cap A\in\mathfrak{E}^{(n)}$, 且 $B\cap A_k,B^c\cap A_k\in\mathfrak{E}^{(n)},k=1,\cdots,m$,

$$A=B\cap A+B^c\cap A,$$

$$B\cap A=B\cap A_1+\cdots+B\cap A_m,$$
$$B^c\cap A=B^c\cap A_1+\cdots+B^c\cap A_m.$$

由于 $B\cap A_2=\varnothing,B^c\cap A_1=\varnothing$, 所以 $B\cap A$ 及 $B^c\cap A$ 都是个数少于 m 的有限区间的和. 利用 $l=2$ 的结论及归纳假设, 得

$$\begin{aligned}\mu(A)&=\mu(B\cap A)+\mu(B^c\cap A)\\&=\mu(B\cap A_1)+\cdots+\mu(B\cap A_m)\\&\quad+\mu(B^c\cap A_1)+\cdots+\mu(B^c\cap A_m)\\&=\mu(A_1)+\cdots+\mu(A_m).\end{aligned}$$

(23) 式获证.

ii) 再证: 当 $[a,b)$ 是无限区间时, 由 (22) 定义的 $\mu([a,b))$ 有意义, 即

$$\lim_{N\to\infty}\mu([a,b)\cap I^{(N)}) \tag{28}$$

存在 (可能是 $+\infty$).

事实上, 由于 $A^{(N)}\triangleq[a,b)\cap I^{(N)}$ 是有限区间, 且 $A^{(N)}\subset A^{(N+1)}$, 由 $\mathfrak{E}^{(n)}$ 是半集代数知有 $\mathfrak{E}^{(n)}$ 中两两不相交的区间 $C_1,\cdots,C_p$ 使

$$A^{(N+1)}=\sum_{r=1}^{p}C_r,\quad C_1=A^{(N)},$$

显然 $C_r\subset A^{(N+1)},r=1,\cdots,p$, 并且都是有限区间, 因而由分布函数的定义及 (21) 知 $\mu(C_r)\geqslant 0$. 再由 i) 知

$$\mu(A^{(N+1)})=\mu(A^{(N)})+\sum_{r=2}^{p}\mu(C_r)\geqslant\mu(A^{(N)}),$$

即 $\mu([a,b)\cap I^{(N)})$ 是不减序列, 故 (28) 式中的极限永远存在.

于是 (21) 及 (22) 在 $\mathfrak{E}^{(n)}$ 上定义了一个非负集函数. 我们指出, 对于有限区间 $[a,b),\mu([a,b))$ 也可以写成 (22) 中的极限形式. 这是由于当 N 充分大时 $[a,b)\subset I^{(N)}$, 因而 $[a,b)\cap I^{(N)}=[a,b)$.

$$\mu([a,b))=\lim_{N\to\infty}\mu([a,b)\cap I^{(N)}),[a,b)\in\mathfrak{E}^{(n)}. \tag{29}$$

这一事实在以下证明中要用到.

iii) 最后证: 由 (21), (22) 在 $\mathfrak{E}^{(n)}$ 上定义的集函数 μ 是有限可加的. 因为在 i) 中已经证明 μ 在 $\mathfrak{E}^{(n)}$ 的一切有限区间类上是有限可加的, 所以若 $A_k \in \mathfrak{E}^{(n)}, k = 1, \cdots, l$ 两两不交, 且 $A = \sum\limits_{k=1}^{l} A_k \in \mathfrak{E}^{(n)}$, 令

$$A^{(N)} = A \cap I^{(N)}, A_k^{(N)} = A_k \cap I^{(N)}, k = 1, \cdots, l,$$

则 $A^{(N)} \in \mathfrak{E}^{(n)}$, 且 $A_k^{(N)} \in \mathfrak{E}^{(n)}, k = 1, \cdots, l$ 是两两不交的有限区间且

$$A^{(N)} = \sum_{k=1}^{l} A_k^{(N)}.$$

于是由 i) 知

$$\mu(A^{(N)}) = \sum_{k=1}^{l} \mu(A_k^{(N)}).$$

而由 (22) 及 (29) 知当 $N \to \infty$ 时左端的极限是 $\mu(A)$, 右端的每一项 $\mu(A_k^{(N)})$ 的极限是 $\mu(A_k)$, 故

$$\mu(A) = \sum_{k=1}^{l} \mu(A_k).$$

引理获证. □

引理 3 若 $F(x_1, \cdots, x_n)$ 是 $R^{(n)}$ 上的分布函数, $a, b \in R^{(n)}, a \leqslant b$, 则

$$\lim_{g \to 0+} \Delta_{b, a-g} F = \Delta_{b,a} F, \tag{30}$$

$$\lim_{g \to 0+} \Delta_{b-g, a} F = \Delta_{b,a} F. \tag{31}$$

证 由于方法类似, 只证 (30) 式.

首先, 由引理 2 知 μ 是 $\mathfrak{E}^{(n)}$ 上的有限可加测度, 因而若 $c^{(i)}, d^{(i)} \in R^{(n)}, i = 1, 2, [c^{(1)}, d^{(1)}) \subset [c^{(2)}, d^{(2)})$, 则 $\Delta_{d^{(1)}, c^{(1)}} F = \mu([c^{(1)}, d^{(1)})) \leqslant \mu([c^{(2)}, a^{(2)})) = \Delta_{d^{(2)}, c^{(2)}} F$. 我们选择 $\varepsilon^{(m)} = \left(\frac{1}{m}, \cdots, \frac{1}{m}\right), m \geqslant 1$, 则由 $[a - \varepsilon^{(m)}, b) \supset [a - \varepsilon^{(m+1)}, b)$ 知 $\Delta_{b, a-\varepsilon^{(m)}} F \geqslant \Delta_{b, a-\varepsilon^{(m+1)}} F$, 当 $m \to \infty$ 时 $\Delta_{b, a-\varepsilon^{(m)}} F$ 是单调下降序列, 故极限存在, 设为 G. 即任给 $\varepsilon > 0$, 存在 m, 使

$$0 \leqslant \Delta_{b, a-\varepsilon^{(m)}} F - G < \varepsilon,$$

设 $g = (g_1, \cdots, g_n), 0 < g_k < \frac{1}{m}, k = 1, \cdots, n$; 必存在 l, 使 $g_k > \frac{1}{m+l}, k = 1, \cdots, n$, 故有

$$G \leqslant \Delta_{b, a-\varepsilon^{(m+l)}} F \leqslant \Delta_{b, a-g} F \leqslant \Delta_{b, a-\varepsilon^{(m)}} F.$$

因此, 当 $0 < g < \varepsilon^{(m)}$ 时

$$0 \leqslant \Delta_{b,a-g}F - G < \varepsilon. \tag{32}$$

此即

$$\lim_{g\to 0+} \Delta_{b,a-g}F = G.$$

今往证 $G = \Delta_{b,a}F$. 在 (32) 式中取累次极限, 逐次令 $g_1 \to 0+, \cdots, g_n \to 0+$, 利用 $F(x_1, \cdots, x_n)$ 对每个自变量的左连续性, 即得

$$\lim_{g_1\to 0+} \cdots \lim_{g_n\to 0+} \Delta_{b,a-g}F = \Delta_{b,a}F.$$

故

$$0 \leqslant \Delta_{b,a}F - G < \varepsilon.$$

由于 ε 的任意性, 故得 $G = \Delta_{b,a}F$. 即 (30) 获证. □

引理 4　若 $F(x_1, \cdots, x_n)$, 是 $R^{(n)}$ 上的分布函数, 则在引理 2 中由 (21) 及 (22) 定义的集函数 μ, 在 $\mathfrak{E}^{(n)}$ 上是 $\sigma-$ 可加的.

证　需要证明: 若 $A_k \in \mathfrak{E}^{(n)}, k = 1, 2, \cdots$, 两两不交且 $A = \sum_{k=1}^{\infty} A_k \in \mathfrak{E}^{(n)}$, 则

$$\mu(A) = \sum_{k=1}^{\infty} \mu(A_k). \tag{33}$$

i) 先证 $\sum_{k=1}^{\infty} \mu(A_k) \leqslant \mu(A)$. 对任一取定的 $n, A_1, \cdots, A_n \subset A, A_k$ 两两不交, μ 是 $\mathfrak{E}^{(n)}$ 上的有限可加测度, 由 §1.5 性质 1 知

$$\sum_{k=1}^{n} \mu(A_k) \leqslant \mu(A).$$

令 $n \to \infty$, 即得 $\sum_{k=1}^{\infty} \mu(A_k) \leqslant \mu(A)$.

ii) 再证 $\sum_{k=1}^{\infty} \mu(A_k) \geqslant \mu(A)$. 不妨设 $A \neq \varnothing$.

先设 A 为有限区间, 于是 $A_k, k = 1, 2, \cdots$ 皆为有限区间. 令

$$A = [c, d), A_k = [c^{(k)}, d^{(k)}), k = 1, 2, \cdots.$$

则由引理 3 知

$$\lim_{g\to 0+} \mu([c^{(k)} - g, d^{(k)}) = \mu([c^{(k)}, d^{(k)})),$$

$$\lim_{g\to 0+}\mu([c,d-g))=\mu([c,d)).$$

任给 $\varepsilon>0$, 存在 $g>0$(指 g 的每一分量大于零), 使

$$\mu(A)-\varepsilon<\mu([c,d-g)), \tag{34}$$

对每一 $k\geqslant 1$, 存在 $g^{(k)}>0$, 使

$$\mu([c^{(k)}-g^{(k)},d^{(k)}))\leqslant\mu(A_k)+\frac{\varepsilon}{2^k}. \tag{35}$$

于是

$$[c,d-g]\subset A=\sum_{k=1}^{\infty}A_k\subset\bigcup_{k=1}^{\infty}(c^{(k)}-g^{(k)},d^{(k)}).$$

应用 $R^{(n)}$ 中的有限覆盖定理知存在 N, 使

$$[c,d-g]\subset\bigcup_{k=1}^{N}(c^{(k)}-g^{(k)},d^{(k)}).$$

于是

$$[c,d-g]\subset\bigcup_{k=1}^{N}[c^{(k)}-g^{(k)},d^{(k)}).$$

应用半集代数上有限可加测度的半可加性 (§1.5 性质 2) 及 (35) 式知

$$\mu([c,d-g))\leqslant\sum_{k=1}^{N}\mu([c^{(k)}-g^{(k)},d^{(k)}))$$

$$\leqslant\sum_{k=1}^{N}\left[\mu(A_k)+\frac{\varepsilon}{2^k}\right]<\sum_{k=1}^{\infty}\mu(A_k)+\varepsilon.$$

再由 (34) 式知

$$\mu(A)-\varepsilon\leqslant\mu([c,d-g))<\sum_{k=1}^{\infty}\mu(A_k)+\varepsilon.$$

由于 ε 的任意性, 即得

$$\mu(A)\leqslant\sum_{k=1}^{\infty}\mu(A_k).$$

再设 A 是无限区间. 令 $A^{(N)}=A\cap I^{(N)}, A_k^{(N)}=A_k\cap I^{(N)}, k=1,2,\cdots$, 两两不交, 则 $A^{(N)}=\sum_{k=1}^{\infty}A_k^{(N)}$, 由已证结果及 $\mu(A_k^{(N)})\leqslant\mu(A_k)$ 知

$$\mu(A^{(N)})\leqslant\sum_{k=1}^{\infty}\mu(A_k^{(N)})\leqslant\sum_{k=1}^{\infty}\mu(A_k).$$

令 $N \to \infty$, 即得

$$\mu(A) \leqslant \sum_{k=1}^{\infty} \mu(A_k).$$

于此, 引理 4 获证. □

由以上四个引理可知由 (21) 及 (22) 定义的 μ 在半集代数 $\mathfrak{E}^{(n)}$ 上是测度, 因为它在有限区间上取有限值, 因而是 $\sigma-$ 有限的. 由 §1.5 定理 2 知在 $\mathscr{B}^{(n)} = \sigma(\mathfrak{E}^{(n)})$ 上存在唯一的测度 μ_F 满足 (20). 再由 §1.5 定理 8 知它可以唯一决定在 μ_F^* 可测集类上的测度, 即 L–S 测度. 至此定理 2 获证. □

下面我们来证明定理 1 的逆定理.

定理 3 设 $F(x_1, \cdots, x_n)$ 是定义在 $R^{(n)}$ 上的有限值函数, 且满足

(a) $\Delta_{b,a}F \geqslant 0$ 对一切 $a, b \in R^{(n)}, a \leqslant b$.

(b) $F(x_1, \cdots, x_n)$ 对每一自变量左连续.

(c) 对于任意的 $1 \leqslant l \leqslant n$, 任意的 $1 \leqslant k_1 < \cdots < k_1 \leqslant n, F(\cdots, x_{k_1-1}, -\infty, x_{k_1+1}, \cdots, x_{k_l-1}, -\infty, x_{k_l+1}, \cdots) = 0$.

(d) $F(\infty, \cdots, \infty) = 1$.

则 $F(x_1, \cdots, x_n)$ 为某一 n 维随机变量 $(\xi_1, \cdots, \xi_n)$ 的分布函数. 称满足上述条件的分布函数为概率分布函数.

证 i) 由于 $F(x_1, \cdots, x_n)$ 满足 (a), (b), 因而是 $R^{(n)}$ 上分布函数, 由定理 2 知在 $\mathscr{B}^{(n)}$ 上存在唯一的测度 μ_F 满足

$$\mu_F([a, b)) = \Delta_{b,a}F, \text{对一切} a, b \in R^{(n)}, a \leqslant b.$$

由 (c) 可知, 对一切 $b \in R^{(n)}$,

$$\begin{aligned}\mu_F((-\infty, b)) &= \lim_{N \to \infty} \mu_F((-\infty, b) \cap I^{(N)}) = \lim_{N \to \infty} \Delta_{b,-N}F \\ &= F(b_1, \cdots, b_n).\end{aligned} \tag{36}$$

再由 (d) 及测度的下连续性可知

$$\begin{aligned}\mu_F((-\infty, \infty)) &= \lim_{N \to \infty} \mu_F((-\infty, N)) = \lim_{N \to \infty} F(N, \cdots, N) \\ &= F(\infty, \cdots, \infty) = 1.\end{aligned} \tag{37}$$

故 μ_F 是 $R^{(n)}$ 上的概率测度, $(R^{(n)}, \mathscr{B}^{(n)}, \mu_F)$ 是概率场.

ii) 在概率场 $(R^{(n)}, \mathscr{B}^{(n)}, \mu_F)$ 上定义 n 个随机变量:

$$\xi_k(x_1, \cdots, x_n) = x_k, (x_1, \cdots, x_n) \in R^{(n)}, k = 1, \cdots, n. \tag{38}$$

则 $\xi_1, \cdots, \xi_n$ 显然是 $(R^{(n)}, \mathscr{B}^{(n)})$ 上取有限值的可测函数, 因而是随机变量. 由 (36) 知其分布函数为

$$\mu_F\left(\bigcap_{k=1}^{n}\{\xi_k < x_k\}\right) = \mu_F((-\infty, x)) = F(x_1, \cdots, x_n).$$

定理 2 的另一个特例就是 n 维实空间上的 Lebesgue 测度. 所谓 Lebesgue 测度就是在定理 2 中令 $F(x_1, \cdots, x_n) = x_1 \cdot x_2 \cdots x_n$ (显然它是一个分布函数) 而建立起来的测度 μ, 它在 $\sigma(\mathfrak{E}^{(n)}) = \mathscr{B}^{(n)}$ 的元 A 上的值为 A 的体积. 事实上它在一切可求积的区域 (按数学分析中的定义) 上都有定义, 因为可求积区域一定是 μ_F^* 可测集. 这些事实留给读者证明.

2.3.3 随机变量概率规律的全面描述 —— 分布律的概念及其与分布函数的关系

本小节目的在于说明如何完全地描述随机变量的概率规律, 从而更充分地说明分布函数的作用. §2.1 指出要想了解随机变量的规律, 就是要知道随机变量取这些值或那些值的概率. 首先要求知道 ξ 取每一区间内的点的概率, 例如 $\xi \in [a, b), \xi \in (a, b), \xi \in (a, b], \xi \in [a, b]$(其中 $\xi = (\xi_1, \cdots, \xi_N), a = (a_1, \cdots, a_n), b = (b_1, \cdots, b_n)$) 以及 ξ 取其他各种类型的区间内的点的概率. 例如, 设 (ξ, η) 表示向目标 (a, b) 进行射击时弹丸着点的位置, 那么就需要知道弹丸着点距目标之差不大于 ε, 即 $(\xi - a)^2 + (\eta - b)^2 \leqslant \varepsilon^2$ 的概率, 此即 $(\xi, \eta) \in G, G = \{(x, y) : (x - a)^2 + (y - b)^2 \leqslant \varepsilon^2\}$ 的概率. 诸如此类的问题, 归纳起来就是需要知道由

$$P_\xi(B) = P(\xi \in B), \quad B \in \mathscr{B}^{(n)} \tag{39}$$

确定的 $R^{(n)}$ 中一切 Borel 集上定义的集函数 P_ξ. 因此可以认为集函数 P_ξ 完全描述了 ξ 的概率规律.

定义 2 设 ξ 是 $(\Omega, \mathscr{A}, P)$ 上的一个 n 维实随机变量, 称由 (39) 确定的 $\mathscr{B}^{(n)}$ 上的非负集函数 P_ξ 为 ξ 的概率分布.

事实上, P_ξ 是 $(R^{(n)}, \mathscr{B}^{(n)})$ 上的概率测度.

在理论上, 一般我们不去研究 P_ξ, 而是通过研究分布函数来掌握随机变量的概率特征. 因为由定理 2 知 P_ξ 是由 F 唯一确定的. 再者, 分布函数是一种特殊的点函数, 而点函数在古典分析中已有了大量的研究, 它便于计算, 因此分布函数在概率论中有着重要的理论意义及实际意义. 我们有

定义 3 随机变量 ξ 的能唯一决定 ξ 的概率分布 P_ξ 的任何规律称为 **ξ的分布律**.

显然 ξ 的分布函数是它的分布律. 通常我们并不考虑随机变量是定义在什么概率场上的函数, 而是通过定理 3 规定出一个概率场 $(R^{(n)}, \mathscr{B}^{(n)}, P)$, 在其上用坐

标函数定义随机变量, 即 $\xi_k(x_1,\cdots,x_n)=x_k$, 这时, $\xi=(\xi_1,\cdots,\xi_n)$ 的分布函数就是预先给定的那一个. 于是, 对具体的随机变量, 只要给了分布函数, 就将它的研究归结到一般理论的基础上了. 但需要注意给定的分布函数必须满足定理 3 的条件, 即必须是一个概率分布函数.

在 §2.1 例 8、例 9 中我们直接给出了 ξ 的分布函数, 必须验证所给出的函数确实是概率分布函数.

对于一维正态分布 $N(m,\sigma^2)$,

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-m)^2}{2\sigma^2}}dt.$$

显然 $\Delta_{b,a}F\geqslant 0$(当 $a\leqslant b$ 时), 且 $F(x)$ 是连续的. 因而定理 3 中条件 a), b) 满足, 再由于积分是收敛的, 可知

$$\lim_{x\to-\infty}F(x)=0.$$

因而条件 c) 满足, 为了验证条件 d) 满足, 令 $t_1=\dfrac{t-m}{\sqrt{2}\sigma}$, 则 $dt=\sqrt{2}\sigma dt_1$, 而

$$\begin{aligned}\lim_{x\to\infty}F(x)=&\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{-\frac{(t-m)^2}{2\sigma^2}}dt\\ =&\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}e^{-t_1^2}dt_1=\frac{2}{\sqrt{\pi}}\int_{0}^{\infty}e^{-t_1^2}dt_1=1.\end{aligned}$$

故 d) 满足, 因而 $F(x)$ 确实是概率分布函数.

对于 n 维正态分布 $N(m,D),m=(m_1,\cdots,m_n),D$ 为 n 阶正定方阵, 我们要证 §2.1 例 9 中的函数 F 是概率分布函数. 我们宁愿把问题提得更广一些, 即考虑以下的问题: 设 $Q(x_1,\cdots,x_n)$ 是一正定二次型, 问 n 元函数

$$F(x_1,\cdots,x_n)=c\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}e^{-Q(t_1-m_1,\cdots,t_n-m_n)}dt_1\cdots,dt_n \tag{40}$$

在什么条件下是一概率分布函数. 首先容易看出 $F(x_1,\cdots,x_n)$ 是连续函数, 且对任意 $a,b\in R^{(n)},a\leqslant b$,

$$\Delta_{b,a}F=c\int_{a_1}^{b_1}\cdots\int_{a_n}^{b_n}e^{-Q(t_1-m_1,\cdots,t_n-m_n)}dt_1,\cdots,dt_n,$$

故只要 $c\geqslant 0$, 即可保证 $\Delta_{b,a}F\geqslant 0$. 可知定理 3 中的 a), b) 成立. 只需寻求使定理 3 c), d) 满足的条件.

由于 $Q(x_1,\cdots,x_n)$ 为正定二次型, 故可写成

$$Q(x_1,\cdots,x_n)=XAX',$$

其中 $X=(x_1,\cdots,x_n),X'$ 为 X 的转置, A 为正定方阵. 由线性代数知有一正交方阵 U(即满足 $UU'=I$), 使

$$A=U\Lambda U',$$

其中 Λ 为 $n\times n$ 对角形方阵, 即

$$\Lambda=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix},$$

且 $\lambda_k>0,k=1,\cdots,n$. 于是

$$F(\infty,\cdots,\infty)=c\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}$$

$$\exp\{-(t-m)U\Lambda U'(t-m)'\}dt_1\cdots dt_n.$$

作变量替换 $v=(v_1,\cdots,v_n)=(t-m)U$, 则 $t-m=vU'$, 容易验证这个变换的函数行列式

$$\frac{\partial(t_1,\cdots,t_n)}{\partial(v_1,\cdots,v_n)}=|U|=\pm1,$$

其中 $|U|$ 表示 U 的行列式, 故

$$\begin{aligned}F(\infty,\cdots,\infty)&=c\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\{-v\Lambda v'\}dv_1\cdots dv_n\\&=c\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left\{-\sum_{k=1}^{n}\lambda_k v_k^2\right\}dv_1\cdots,dv_n\\&=c\prod_{k=1}^{n}\left(\int_{-\infty}^{\infty}e^{-\lambda_k v_k^2}dv_k\right)=c\prod_{k=1}^{n}\sqrt{\frac{\pi}{\lambda_k}}.\end{aligned}$$

故欲使 $F(\infty,\cdots,\infty)=1$, 必须

$$c=\frac{(\lambda_1\cdots\lambda_n)^{\frac{1}{2}}}{\pi^{\frac{n}{2}}}.$$

且知由 (40) 定义的广义积分是收敛的, 故定理 3 c) 满足. 又因为 $|A|=|\Lambda||U|^2=|\Lambda|$,

$$(\lambda_1,\cdots\lambda_n)^{\frac{1}{2}}=\sqrt{|A|},$$

故当

$$c=\frac{\sqrt{|A|}}{(\pi)^{\frac{n}{2}}}.$$

时, 由 (40) 定义的函数 $F(x_1,\cdots,x_n)$ 满足定理 3 的全部条件.

在例 9 中, $A=\frac{1}{2}D^{-1}$, 故 $|A|=\frac{1}{2^n}|D^{-1}|=\frac{1}{2^n|D|}$,

故

$$\frac{1}{(2\pi)^{\frac{n}{2}}|D|^{\frac{1}{2}}}\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}\exp\left\{-\frac{1}{2}\times(t-m)D^{-1}(t-m)'\right\}dt_1\cdots dt_n$$

满足定理 3 的全部条件, 是一个 n 维随机变量的分布函数.

这里我们再给出一般均匀分布的分布函数. 设 D 是 $L-$ 测度有限的集, μ 是 $L-$ 测度, 令

$$F(x_1,\cdots,x_n)=\frac{\mu(D\cap(-\infty,x))}{\mu(D)},\quad (x_1,\cdots,x_n)\in R^{(n)},$$

其中 $x=(x_1,\cdots,x_n),-\infty=(-\infty,\cdots,-\infty)$. 则称 F 为在 D 上均匀分布的随机变量的分布函数. 易证它具有概率分布函数的一切性质.

不仅 ξ 的分布函数是 ξ 的分布律, 而且对离散型随机变量 (即取有限值的简单函数或初等函数). 例如 $P(\xi=x^{(k)})=p_k, x^{(k)}\in R^{(n)}, k=1,2,\cdots,\sum\limits_k p_k=1$, 则对任意 $B\in\mathscr{B}^{(n)}$,

$$P_\xi(B)=P(\xi\in B)=\sum_{x^{(k)}\in B}p_k.$$

$\{P(\xi=x^{(k)})=p_k,k=1,2,\cdots\}$ 称为 ξ 的分布列, 它完全决定了 P_ξ, 因此对离散型随机变量来说, ξ 的分布列也是 ξ 的分布律; 对连续型随机变量 ξ, 存在非负可测函数 $p(x_1,\cdots,x_n)\geqslant 0$, 对一切 $(x_1,\cdots,x_n)\in R^{(n)}$, 且 $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(x_1,\cdots,x_n)$ $dx_1\cdots dx_n=1$, 使 ξ 的分布函数

$$F(x_1,\cdots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}p(t_1,\cdots,t_n)dt_1\cdots,dt_n,$$

对一切 $(x_1,\cdots,x_n)\in R^{(n)}$ 成立, 其中积分暂且理解为 Riemann 积分. $p(t_1,\cdots,t_n),(t_1,\cdots,t_n)\in R^{(n)}$ 称为 ξ 的分布密度, 它完全决定了 P_ξ. 故对连续型随机变量来说, ξ 的分布密度也是 ξ 的分布律.

不仅如此, ξ 的概率分布还可以决定 ξ 的 Borel 函数 (如果有限) 的分布律. 由 §2.2 定理 4 我们有

定理 4 设 $\xi=(\xi_1,\cdots,\xi_n)$ 是 $(\Omega,\mathscr{A},P)$ 上的 n 维随机变量, $f_k(x_1,\cdots,x_n)$, $(x_1,\cdots,x_n)\in R^{(n)},k=1,\cdots,m$, 是实可测函数, 且 $P(f_k(\xi)=\pm\infty)=0,k=1,\cdots,m$. 则 $(f_1(\xi),\cdots,f_m(\xi))$ 是一个 m 维实随机变量. 它的分布函数

$$F(y_1,\cdots,y_m)=P_\xi\left(\bigcap_{k=1}^{m}f_k^{-1}(-\infty,y_k)\right),$$

其中 P_ξ 为 ξ 的概率分布.

习题及补充

1. 设定理 2 中的分布函数 F 是 $R^{(n)}$ 上的连续函数, μ 为由 F 决定的 L–S 测度, 试证

i) $\mu([a,b]) = \mu((a,b)) = \mu((a,b]) = \mu([a,b))$;

ii) $R^{(n)}$ 中任一可数集的测度是零.

2. 设 F,G 是 $R^{(n)}$ 上的分布函数, 试证: 由 F 和 G 决定的 L–S 测度 $\mu_F = \mu_G$ 的充分必要条件是

$$
\begin{aligned}
&F(x_1,\cdots,x_n) - G(x_1,\cdots,x_n) \\
=&\sum_{i=1}^{N} f_i(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n), (x_1,\cdots,x_n) \in R^{(n)},
\end{aligned}
$$

其中 $f_i(x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n), i=1,\cdots,n$, 是 $n-1$ 元函数.

3. 设 $(\xi_1,\cdots,\xi_n)$ 是 n 维离散型随机变量, 而其分布函数是 $F(x_1,\cdots,x_n), \xi_k$ 只取 $x_{kl_k}, l_k = 1,2,\cdots$ 为值, 试用 F 表出 $P(\xi_1 = x_{1l_1},\cdots,\xi_n = x_{nl_n})$.

4. 设 $F(x)$ 是概率分布函数, 试证对任何 $h \neq 0$, 函数

$$
\varPhi_1(x) = \frac{1}{h}\int_x^{x+h} F(t)dt,
$$

$$
\varPhi_2(x) = \frac{1}{2h}\int_{x-h}^{x+h} F(t)dt,
$$

都是概率分布函数.

5. 设 $\{r_1,r_2,\cdots\}$ 是有理数全体, 令

$$
F(x) = \sum_{r_k<x} \frac{1}{2^k}.
$$

试证: $F(x)$ 是概率分布函数, 其不连续点集为 $\{r_1,r_2,\cdots\}$.

注 此题给出一个不连续点在 $R^{(1)}$ 上稠密的离散型分布函数的例子.

6. 已知 μ 是 $R^{(n)}$ 上的 L–S 测度, 试构造一个分布函数 F, 使由 F 确定的测度就是 μ. (对 $n=2$ 的情形加以证明.)

§2.4 独立随机变量

现在我们讨论随机变量的独立性. 随机变量独立的概念是概率论中最重要的概念之一.

我们用一个字母表示向量, 例如: 用 ξ 表示 $(\xi_1,\cdots,\xi_n)$, x 表示 $(x_1,\cdots,x_n)$, $\xi<x$ 表示 $\xi_k<x_k, k=1,\cdots,n$.

定义 1 设 $\xi^{(t)}=(\xi_{t1},\cdots,\xi_{tm_t}), t\in T$ 是概率场 $(\Omega,\mathscr{A},P)$ 上的实随机变量族, T 是任意指标集. 若对任意的正整数 l, $\{t_1,\cdots,t_l\}\subset T$及$x^{(t_s)}\in R^{(m_{t_s})}, s=1,\cdots,l$ 来说, 等式

$$P(\xi^{(t_1)}<x^{(t_1)},\cdots,\xi^{(t_l)}<x^{(t_l)})=\prod_{s=1}^{l}P(\xi^{(t_s)}<x^{(t_s)}) \tag{1}$$

成立, 则称 $\xi^{(t)}, t\in T$ **为独立的随机变量族**.若 T 是可数集, 即 $T=\{1,2,\cdots\}$, 则称 $\{\xi^{(k)}\}$ **为独立随机变量序列.**

若 $\xi_{t_s}=\eta_{t_s}+i\zeta_{t_s}, s=1,\cdots,m_t$ 是复随机变量, 则当 $\xi^{(t)}=(\eta_{t1},\cdots,\eta_{tm_t},\zeta_{t1},\cdots,\zeta_{tm_t}), t\in T$ 独立时, 称 $\xi^{(t)}, t\in T$ 为独立的随机变量族.

性质 1 随机变量族 $\xi^{(t)}, t\in T$ 独立的充分必要条件是它的任何一个有限子集中的随机变量独立.

性质 2 设随机变量族 $\xi^{(t)}, t\in T$ 独立, 若将它分成一些不相交的有限子集, 然后将每一有限子集中的随机变量的分量合起来组成新的随机变量, 则得到的新的随机变量族仍然独立.

由于性质 1, 在本节余下部分只讨论有限个随机变量独立的性质.

应用 §1.6 独立类扩张定理得到

定理 1 设 $\xi^{(k)}, k=1,\cdots,n$ 是 n 个实 (复) 随机变量, 则 $\xi^{(k)}, k=1,\cdots,n$ 独立的充分必要条件是对任何 $B^{(m_k)}\in\mathscr{B}^{(m_k)}(\mathscr{B}_z^{(m_k)}), k=1,\cdots,n$, 有

$$P\left(\bigcap_{k=1}^{n}\{\xi^{(k)}\in B^{(m_k)}\}\right)=\prod_{k=1}^{n}P(\xi^{(k)}\in B^{(m_k)}). \tag{2}$$

证 条件的充分性显然, 今证必要性. 设 $\xi^{(1)},\cdots,\xi^{(n)}$ 是实 (复) 独立随机变量, 则由定义 1 知随机事件类

$$\mathfrak{E}_k=\{\{\xi^{(k)}<x^{(k)}\}:x^{(k)}\in R^{(m_k)}\}$$

(相应地: $\mathfrak{E}_k=\{\{\xi^{(k)}<x^{(k)}\}:x^{(k)}\in R^{(2m_k)}\}$), $k=1,\cdots,n$ 是独立的. 容易看出这一事件类对有限交封闭, 由独立类扩张定理知 $\sigma(\mathfrak{E}_k), k=1,\cdots,n$ 独立, 再由 §2.2 性质 3 知

$$\sigma(\mathfrak{E}_k)=\{\{\xi^{(k)}\in B^{(mk)}\}:B^{(mk)}\in\mathscr{B}^{(m_k)}\})$$

(相应地: $\sigma(\mathfrak{E}_k)=\{\{\xi^{(k)}\in B^{(m_k)}\}:B^{(m_k)}\in\mathscr{B}_z^{(m_k)}\}), k=1,\cdots,n$.

由独立事件类的定义知 (2) 成立. □

由事件独立的直观含义及定理 1 可知, 随机变量独立的直观含义是: $\xi^{(k)}, k=1,\cdots,n$ 各自取任何范围内的值是概率无关的.

推论 1 设随机变量 $\xi^{(k)}, k=1,\cdots,n$ 独立, $\xi^{(k)'}$ 是 $\xi^{(k)}$ 的部分分量组成的随机变量, 则 $\xi^{(k)'}, k=1,\cdots,n$ 仍然独立.

证 设 m_k' 是 $\xi^{(k)'}$ 的维数, 则事件 $\{\xi^{(k)'}\in B^{(m_k')}\}, B^{(m_k')}\in\mathscr{B}^{(m_k')}$(当 $\xi^{(k)'}$ 为复随机变量时相应地为 $\mathscr{B}_z^{(m_k')}$) 必然等于某一事件 $\{\xi^{(k)}\in B^{(m_k)}\}, B^{(m_k)}\in\mathscr{B}^{(m_k)}$(相应地 $\mathscr{B}_z^{(m_k)}$), 因此由 (2) 知 $\xi^{(k)'}, k=1,\cdots,n$ 独立. □

推论 2 若 $\xi^{(k)}, k=1,\cdots,n$ 是 n 个独立的复随机变量, 则 $\xi^{(k)}$ 的实部作成的随机变量 $\eta^{(k)}=(\eta_{k1},\cdots,\eta_{km_k}), k=1,\cdots,n$ 独立, 而其虚部作成的随机变量 $\zeta^{(k)}=(\zeta_{k1},\cdots,\zeta_{km_k}), k=1,\cdots,n$ 也独立.

推论 3 若 $\xi^{(k)}, k=1,\cdots,n$ 独立, 而 $f_k(x^{(k)}), k=1,\cdots,n$ 是 n 个 Borel 可测函数, 若每一 $f_k(\xi^{(k)})$ 有限, 则随机变量 $f_k(\xi^{(k)}), k=1,\cdots,n$ 独立.

证 由于

$$\{f_k(\xi^{(k)})\in B_k^{(1)}\}=\{\xi^{(k)}\in B_k^{(m_k)}\},$$

其中 $B_k^{(1)}\in\mathscr{B}^{(1)}$(或当 f_k 为复 Borel 函数时 $B_k^{(1)}\in B_z^{(1)}$), 而 $B_k^{(m_k)}=f_k^{-1}(B_k^{(1)})\in\mathscr{B}^{(m_k)}$(相应地 $B_z^{(m_k)}$), 故应用定理 1 即得

$$P\left(\bigcap_{k=1}^n\{f_k(\xi^{(k)})\in B_k^{(1)}\}\right)=P\left(\bigcap_{k=1}^n\{\xi^{(k)}\in B_k^{(m_k)}\}\right)$$
$$=\prod_{k=1}^n P(\xi^{(k)}\in B_k^{(m_k)})=\prod_{k=1}^n P(f_k(\xi^{(k)})\in B_k^{(1)}).$$

故 $f_k(\xi^{(k)}), k=1,\cdots,n$, 为独立随机变量. □

若将推论 3 中的函数 $f_k(x^{(k)})$ 换成向量函数, 则推论 3 的结论仍然成立. 请读者自证.

例 1 若 $\xi^{(1)},\cdots,\xi^{(n)}$ 独立且均匀 m 维随机变量, 则 $\xi^{(1)}+\cdots+\xi^{(n-1)}$ 与 $\xi^{(n)}$ 独立.

若令 $\xi^{(k)}=(\xi_{k_1},\cdots,\xi_{km}), t=(t_1,\cdots,t_m)$, 则 $\exp\left(i\sum_{l=1}^m\xi_{nl}t_l\right)$ 与 $\exp\left(i\sum_{l=1}^m\left(\sum_{k=1}^{n-1}\xi_{kl}\right)t_l\right)$ 独立; $\exp\left(i\sum_{l=1}^m\xi_{kl}t_l\right), k=1,\cdots,n$ 独立.

下面给出随机变量独立的判别条件. 为了书写简单, 用 $(\xi^{(1)},\cdots,\xi^{(n)})$ 表示 $(\xi_{11},\cdots,\xi_{1m_1},\cdots,\xi_{n1},\cdots,\xi_{nm_n})$, 用 $F(x^{(1)},\cdots,x^{(n)})$ 表示它的分布函数. $x^{(k)}\to\infty$ 表示 $x_{kr_k}\to\infty, r_k=1,\cdots,m_k$.

定理 2 设 $\xi^{(k)}, k=1,\cdots,n$ 是 n 个独立的实随机变量, 则 $(\xi^{(1)},\cdots,\xi^{(n)})$ 的分布函数等于 $\xi^{(k)}$ 的分布函数的乘积. 反之, 若 $(\xi^{(1)},\cdots,\xi^{(n)})$ 的分布函数

$$F(x^{(1)},\cdots,x^{(n)})=\prod_{k=1}^{n}F_k(x^{(k)}) \tag{3}$$

且 $F_k(\infty)=1$, 则 $\xi^{(k)}, k=1,\cdots,n$ 独立, 且 $F_k(x^{(k)})$ 是 $\xi^{(k)}$ 的分布函数.

证 定理的第一部分为定义条件 (1) 的直接推论, 今证定理的第二部分.

由 §2.3 性质 4 知, $\xi^{(k)}$ 的分布函数是 $F(\infty,\cdots,\infty,x^{(k)},\infty,\cdots,\infty)$, 而由 (3) 及 $F_k(\infty)=1$ 知

$$F(\infty,\cdots,\infty,x^{(k)},\infty,\cdots,\infty)=F_k(x^{(k)}).$$

故 $\xi^{(k)}$ 的分布函数为 $F_k(x^{(k)})$. 其次, 对任意 $k_1,\cdots,k_l(l\leqslant n)$ 及 $x^{(k_r)}\in R^{m_{k_r}}, r=1,\cdots,l$, 在 (3) 中令 $x^{(k)}\to\infty, k\neq k_r$, 即得

$$\begin{aligned}&P\left(\xi^{(k_1)}<x^{(k_1)},\cdots,\xi^{(k_l)}<x^{(k_l)}\right)\\&=F(\infty,\cdots,\infty,x^{(k_1)},\infty,\cdots,\infty,x^{(k_l)},\infty,\cdots,\infty)\\&=F_{k_1}(x^{(k_1)})\cdots F_{k_l}(x^{(k_l)})=\prod_{r=1}^{l}p(\xi^{(k_r)}<x^{(k_r)}),\end{aligned}$$

此即 (1) 式. 故知 $\xi^{(k)}, k=1,\cdots,n$ 独立. □

我们还可以将定理的第二部分加强一点成为

推论 4 设 $\xi^{(k)}, k=1,\cdots,n$ 是 n 个多维实随机变量, 若有 n 个函数 $G_k(x^{(k)})$ 使得 $(\xi^{(1)},\cdots,\xi^{(n)})$ 的分布函数

$$F(x^{(1)},\cdots,x^{(n)})=\prod_{k=1}^{n}G_k(x^{(k)}), \tag{4}$$

则 $\xi^{(k)}, k=1,\cdots,n$ 独立, 且 $\lim\limits_{x^{(k)}\to\infty}G_k(x^{(k)})=G_k(\infty)$ 存在, $G_k(\infty)\neq 0$ 而 $\xi^{(k)}$ 的分布函数为 $G_k(x^{(k)})/G^{(k)}(\infty)$. 证明留给读者.

这个推论在验证随机变量的独立性时比定理2方便, 因为它不需要验证$G_k(\infty)$=1.

定理 3 设 $\xi^{(k)}, k=1,\cdots,n$ 是离散型 (实或复) 随机变量, 且 $\xi^{(k)}$ 只取 $\{x_s^{(k)}\}(x_s^{(k)}\in R^{(m_k)}$ 或 $Z^{(m_k)})$ 中的值, 令

$$P(\xi^{(k)}=x_s^{(k)})=p_s^{(k)}, s=1,2,\cdots,$$

则 $\xi^{(k)}, k=1,\cdots,n$ 独立的充分必要条件是

$$P(\xi^{(1)}=x_{s_1}^{(1)},\cdots,\xi^{(n)}=x_{s_n}^{(n)})=\prod_{k=1}^{n}p_{s_k}^{(k)} \tag{5}$$

对一切 $s_1, \cdots, s_n$ 成立.

证 i) 若 $\xi^{(k)}, k=1,\cdots,n$ 独立, 由于单元素集是 Borel 集, 故由定理 1 知 (5) 式成立. 条件的必要性获证.

ii) 若 (5) 式成立, 则对任何 $B^{(m_k)} \in \mathscr{B}^{(m_k)}$(或 $\mathscr{B}_z^{(m_k)}$), $k=1,\cdots,n$,

$$\begin{aligned}
P\Big(\bigcap_{k=1}^{n}\{\xi^{(k)} \in B^{(m_k)}\}\Big) &= P\Big(\bigcap_{k=1}^{n}\Big(\sum_{x_{s_k}^{(k)} \in B^{(m_k)}} \{\xi^{(k)} = x_{s_k}^{(k)}\}\Big)\Big) \\
&= P\Big(\sum_{x_{s_1}^{(1)} \in B^{(m_1)}} \cdots \sum_{x_{s_n}^{(n)} \in B^{(m_n)}} \{\xi^{(1)} = x_{s_1}^{(1)}, \cdots, \xi^{(n)} = x_{s_n}^{(n)}\}\Big) \\
&= \sum_{x_{s_1}^{(1)} \in B^{(m_1)}} \cdots \sum_{x_{s_n}^{(n)} \in B^{(m_n)}} P(\xi^{(1)} = x_{s_1}^{(1)}, \cdots, \xi^{(n)} = x_{s_n}^{(n)}) \\
&= \sum_{x_{s_1}^{(1)} \in B^{(m_1)}} \cdots \sum_{x_{s_n}^{(n)} \in B^{(m_n)}} \prod_{k=1}^{n} p_{s_k}^{(k)} \\
&= \prod_{k=1}^{n}\Big(\sum_{x_{s_k}^{(k)} \in B^{(m_k)}} p_{s_k}^{(k)}\Big) = \prod_{k=1}^{n} P(\xi^{(k)} \in B^{(m_k)}).
\end{aligned}$$

故由定理 1 知 $\xi^{(k)}, k=1,\cdots,n$ 独立. 条件的充分性获证. □

例 2 若 $(\xi_1,\cdots,\xi_r)$ 是具有参数 $\lambda_1,\cdots,\lambda_r$ 的 r 维 Poisson 分布的随机变量, 则 $\xi_1,\cdots,\xi_n$ 独立, 且 ξ_k 为具有参数 λ_k 的 Poisson 分布的随机变量.

事实上, 由 §2.1 例 6 知

$$\begin{aligned}
F_{\xi_1,\cdots,\xi_r}(x_1,\cdots,x_r) &= e^{-(\lambda_1+\cdots+\lambda_r)} \sum_{0\leqslant m_1<x_1} \cdots \sum_{0\leqslant m_r<x_r} \frac{\lambda_1^{m_1}\cdots\lambda_r^{m_r}}{m_1!\cdots m_r!} \\
&= \prod_{k=1}^{r}\Big(e^{-\lambda_k}\sum_{0\leqslant m_k<x_k}\frac{\lambda_k}{m_k!}\Big).
\end{aligned}$$

若有一 $x_k \leqslant 0$, 第二等式两端皆理解为零. 令

$$F_k(x_k) = e^{-\lambda_k}\sum_{0\leqslant m_k<x_k}\frac{\lambda_k^{m_k}}{m_k!}.$$

易知

$$\lim_{x_k\to\infty} F_k(x_k) = e^{-\lambda_k}\sum_{m_k=0}^{\infty}\frac{\lambda_k^{m_k}}{m_k!} = e^{-\lambda_k}e^{\lambda_k} = 1.$$

故由定理 2 知 $\xi_1,\cdots,\xi_r$ 独立, 且 ξ_k 的分布函数为 $F_k(x_k), k=1,\cdots,r$. 故 ξ_k 为具有参数 λ_k 的 Poisson 分布的随机变量.

习题及补充

1. 证明例 1.

2. 证明推论 4.

3. 将独立随机变量族的概念按下述定义推广成独立可测映射族.

定义 设 T 是任一指标集, $(\Omega,\mathscr{A},P)$ 是一概率场, 对每一 $t\in T,\xi^{(t)}$ 是 $(\Omega,\mathscr{A})$ 到可测空间 $(X_t,\mathscr{B}_t)$ 的可测映射, 若对任意的正整数 $l,\{t_1,\cdots,t_l\}\subset T$ 及任意的 $B_{t_k}\in\mathscr{B}_{t_k},k=1,\cdots,l$ 都有

$$P\left(\bigcap_{k=1}^{l}\{\xi^{(t_k)}\in B_{t_k}\}\right)=\prod_{k=1}^{l}P(\xi^{(t_k)}\in B_{t_k}),$$

则称 $\{\xi^{(t)},t\in T\}$ 为独立可测映射族.

i) 试就每一 $\xi^{(t)},t\in T$ 都是随机变量的情形说明此定义与定义 1 的关系.

ii) 试推广推论 3 的结论并加以证明.

4. 若 $\xi^{(k)},k=1,\cdots,n$ 独立, $\xi^{(k)}$ 是 m_k 维随机变量, $f^{(k)}$ 是 R^{m_k} 到 R^{n_k} 的 Borel 可测向量函数, 即

$$f^{(k)}=(f_1^{(k)},f_2^{(k)},\cdots,f_{n_k}^{(k)}),$$

试证 $f^{(k)}(\xi^{(k)}),k=1,\cdots,n$ 仍独立.

§2.5 随机变量序列的收敛性

本节的目的是介绍随机变量序列的几乎必然收敛、概率收敛和分布律收敛的定义、基本性质以及它们之间的关系. 而随机变量序列的几乎必然收敛、概率收敛分别是一般测度空间上可测函数序列的几乎处处收敛、测度收敛的特殊情形. 所以我们在一般测度空间上讨论可测函数序列的这两种收敛性.

2.5.1 几乎处处收敛

设 $(\Omega,\mathscr{A},\mu)$ 是一个测度空间, 如果一个定义或一种关系除掉一个测度为零的集 N 外处处成立, 我们就称为**关于 μ 几乎处处成立**, 记作 μ - a.e. 或 a.e., 称 N **为例外集**.

这一节我们约定所有的可测函数都是一个任意给定的测度空间 $(\Omega,\mathscr{A},\mu)$ 上的几乎处处有限的可测函数. 所有 “几乎处处” 都理解为关于 μ 几乎处处, 所有的 a.e. 都理解为 μ-a.e.

定义 1 设 $\{f_n\}$ 是一可测函数序列, f 是可测函数, 若存在 $N \in \mathscr{A}, \mu(N)=0$, 使得对于每一 $\omega \in N^c$ 有

$$f_n(\omega) \to f(\omega), \quad (n \to \infty)$$

则称 $\{f_n\}$**几乎处处收敛于** f, 记作

$$f_n \xrightarrow{\text{a.e.}} f.$$

如果对于每一 $\omega \in N^c$,

$$f_n(\omega) - f_m(\omega) \to 0, \quad (m, n \to \infty)$$

则称 $\{f_n\}$ 几乎处处相互收敛, 记作

$$f_m - f_n \xrightarrow{\text{a.e.}} 0.$$

由定义容易看出, $\{f_n\}$ 几乎处处相互收敛, 也可以利用下述形式来叙述: 对每一 $\omega \in N^c$, 对正整数 ν 一致地有

$$f_{n+\nu}(\omega) - f_n(\omega) \to 0 (n \to \infty),$$

或记作

$$f_{n+\nu} - f_n \xrightarrow{\text{a.e.}} 0, (\text{对}\nu\text{一致})^{①}.$$

性质 1 若 $f_n \xrightarrow{\text{a.e.}} f, \{f_{n'}\}$ 是 $\{f_n\}$ 的任一子序列, 则 $f_{n'} \xrightarrow{\text{a.e.}} f$.

性质 2 若 $f_n \xrightarrow{\text{a.e.}} f, f_n \xrightarrow{\text{a.e.}} f'$, 则 $f = f'$, a.e. 这就是说, 几乎处处收敛的极限函数是几乎唯一的.

证 由假定, 存在 $N, N' \in \mathscr{A}, \mu(N) = \mu(N') = 0$, 使

$$f_n(\omega) \to f(\omega), \quad \text{当}\omega \in M^c, (n \to \infty),$$

$$f_n(\omega) \to f'(\omega), \quad \text{当}\omega \in N'^c, (n \to \infty).$$

因此, 利用数序列极限的唯一性, 当 $\omega \in N^c \cap N'^c$ 时, $f(\omega) = f'(\omega)$, 即 $\{f \neq f'\} \subset N \cup N'$, 但

$$\mu(f \neq f') \leqslant \mu(N \cup N') = 0,$$

即 $f = f'$,a.e.

性质 3 若 $f_n \xrightarrow{\text{a.e.}} f, g_n = f_n$, a.e., $g = f$, a.e., 则 $g_n \xrightarrow{\text{a.e.}} g$.

---

① 为了简单起见, 今后在不致混淆的情况下, 将 "对 ν 一致" 省略!

证　由 $f_n \xrightarrow{\text{a.e.}} f$ 知存在 $M \in \mathscr{A}, \mu(M) = 0$, 使

$$f_n(\omega) \to f(\omega), \text{当}\omega \in M^c, n \to \infty\text{时}.$$

设 $M_0 = \{f \neq g\}, M_n = \{f_n \neq g_n\}$, 则 $M_n \in \mathscr{A}$ 且 $\mu(M_n) = 0, n = 0, 1, 2, \cdots$. 令 $N = M \cup \left(\bigcup_{n=0}^{\infty} M_n\right)$, 则 $N \in \mathscr{A}$, 且 $\mu(N) = 0$. 对于每一 $\omega \in N^c = M^c \bigcap \left(\bigcap_{n=0}^{\infty} M_n^c\right)$,

$$g_n(\omega) = f_n(\omega) \to f(\omega) = g(\omega), (n \to \infty)$$

即 $g_n \xrightarrow{\text{a.e.}} g$.

如果我们在一切 $(\Omega, \mathscr{A}, \mu)$ 可测函数类上定义一个等价关系 ①, 即若 $f = f'$,a.e., 则 $f \sim f'$, 按这种等价关系将可测函数分类, 则由性质 2, 3 知几乎处处收敛性是等价类序列向一等价类的收敛性. 换言之, 一提到 a.e. 收敛性, 则此序列中各个函数及其极限函数均可在几乎相等的条件下任意改变, 特别可以把一个 a.e. 有限的函数换成有限函数. (我们还特别指出可以把 a.e. 可测函数 f—— 即存在零测集 N 使 $f\chi_{N^c}$ 可测, 换成一个可测函数.) 这样的改换不会破坏几乎处处收敛性.

性质 4　若 $f_k^{(n)} \xrightarrow{\text{a.e.}} f_k (k = 1, \cdots, m), G(x_1, \cdots, x_m)$ 是 $R^{(m)}$(或 $Z^{(m)}$) 上的连续函数, 则

$$G(f_1^{(n)}, \cdots, f_m^{(n)}) \xrightarrow{\text{a.e.}} G(f_1, \cdots, f_m).$$

特别若 $f_n \xrightarrow{\text{a.e.}} f, g_n \xrightarrow{\text{a.e.}} g$, 则

$$f_n \pm g_n \xrightarrow{\text{a.e.}} f \pm g;$$

$$f_n g_n \xrightarrow{\text{a.e.}} fg;$$

$$cf_n \xrightarrow{\text{a.e.}} cf (c\text{是任意有限常数});$$

$$|f_n|^r \xrightarrow{\text{a.e.}} |f|^r.$$

进一步若 $\{g_n = 0\}, \{g = 0\}$ 皆为零测集, 令 $g_n' = g_n \chi_{(g_n \neq 0)} + \chi_{(g_n = 0)}, g' = g\chi_{(g \neq 0)} + \chi_{(g=0)}$, 则

$$f_n / g_n' \xrightarrow{\text{a.e.}} f / g'.$$

利用数序列收敛的 Cauchy 准则, 得到下面的 a.e. 收敛准则 (特别注意在这一节中我们限定所有的函数都是 a.e. 有限的).

① 若集 S 的某些元素间建立一个关系 "∼" 它具有性质: 反身性 (对任何 $x \in S, x \sim x$)、传递性 (若 $x \sim y, y \sim z$, 则 $x \sim z$) 及对称性 (若 $x \sim y$, 则 $y \sim x$), 则称此关系为等价关系 (参看 [23]).

定理 1 $f_n \xrightarrow{\text{a.e.}}$(某一)$f(n\to\infty)$ 的充分必要条件是 $f_m - f_n \xrightarrow{\text{a.e.}} 0(m,n\to\infty)$.

证 利用性质 3, 不妨假设 f_n, f 皆有限, 则存在 $N\in\mathscr{A}, \mu(N)=0$, 当 $\omega\in N^c$ 时 $f_n(\omega)\to f(\omega)$, 利用 Cauchy 准则, $f_m(\omega)-f_n(\omega)\to 0(m,n\to\infty)$, 即 $f_m-f_n \xrightarrow{\text{a.e.}} 0$.

反之, 若 $f_m - f_n \xrightarrow{\text{a.e.}} 0$, 则存在一 $M\in\mathscr{A}, \mu(M)=0$, 当 $\omega\in M^c$ 时, $f_m(\omega)-f_n(\omega)\to 0(m,n\to\infty)$, 根据 Cauchy 准则, 存在一个有限数 a_ω, 使

$$f_n(\omega)\to a_\omega,\quad (n\to\infty).$$

令

$$f_n(\omega)=\begin{cases} a_\omega, & \omega\in M^c,\\ 0, & \omega\in M.\end{cases}$$

则

$$f_n \xrightarrow{\text{a.e.}} f.$$

且 f 是一有限可测函数. f 的可测性是由于 $\{f_n\chi_{M^c}\}$ 是可测函数序列, 且对一切 $\omega\in,\Omega$ 有

$$f_n(\omega)\chi_{M^c}(\omega)\to f(\omega),\quad (n\to\infty),$$

即 f 是可测函数序列的极限, 因而可测. □

有限值可测函数序列几乎处处收敛到有限值可测函数是我们最关心的问题. 为了给出它的判别条件, 我们首先考虑可测函数序列的收敛域. 以下假定 f_n, f 均为有限可测函数.

根据数列极限的定义, 容易得出下列收敛域的表示式:

$$\begin{aligned}\{f_n\to f\} &= \bigcap_{\varepsilon>0}\bigcup_{n=1}^{\infty}\bigcap_{\nu=1}^{\infty}\{|f_{n+\nu}-f|<\varepsilon\}\\ &= \bigcap_{k=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{\nu=1}^{\infty}\{|f_{n+\nu}-f|<\varepsilon_k\},\end{aligned}\tag{1}$$

其中 $\varepsilon_k\to 0(k\to\infty)$. 由后一等式看出, $\{f_n\}$ 的收敛域 $\{f_n\to f\}$ 是由可数个可测集 $\{|f_{n+\nu-f}|<\varepsilon_k\}$ 通过交和并的运算得出的, 因而是可测集.

利用 $\{f_n\not\to f\}=\{f_n\to f\}^c$, 立刻得出

$$\begin{aligned}\{f_n\not\to f\} &= \bigcup_{\varepsilon>0}\bigcap_{n=1}^{\infty}\bigcup_{\nu=1}^{\infty}\{|f_{n+\nu}-f|\geqslant\varepsilon\}\\ &= \bigcup_{k}\bigcap_{n}\bigcup_{\nu}\{|f_{n+\nu}-f|\geqslant\varepsilon_k\}.\end{aligned}\tag{1$'$}$$

类似地, 可以得出下列二式;

$$\{f_{n+\nu}-f_n\to 0\}=\bigcap_{\varepsilon>0}\bigcup_n\bigcap_\nu\{|f_{n+\nu}-f_n|<\varepsilon\}$$
$$=\bigcap_k\bigcup_n\bigcap_\nu\{|f_{n+\nu}-f_n|<\varepsilon_k\};\tag{2}$$

$$\{f_{n+\nu}-f_n\not\to 0\}=\bigcup_{\varepsilon>0}\bigcap_n\bigcup_\nu\{|f_{n+\nu}-f_n|\geqslant\varepsilon\}$$
$$=\bigcup_k\bigcap_n\bigcup_\nu\{|f_{n+\nu}-f_n|\geqslant\varepsilon_k\}.\tag{2)'$$

由此可见, $\{f_n\to f\},\{f_n\not\to f\},\{f_{n+\nu}-f_n\to 0\},\{f_{n+\nu}-f_n\not\to 0\}$ 都是可测集, 因而 $f_n\xrightarrow{\text{a.e.}}f$ 的充分必要条件是

$$\mu(f_n\not\to f)=0,(n\to\infty).$$

而 $f_{n+\nu}-f_n\xrightarrow{\text{a.e.}}0$ 的充分必要条件是

$$\mu(f_{n+\nu}-f_n\not\to 0)=0,\quad(n\to\infty).$$

更进一步由此可得出下面关于有限可测函数序列几乎处处收敛到有限可测函数及有限可测函数序列几乎处处相互收敛的判别条件.

定理 2 设 $f,f_n,n\geqslant 1$ 都是有限可测函数.

i) $f_n\xrightarrow{\text{a.e.}}f$ 当且仅当对于任一 $\varepsilon>0$,

$$\mu\left(\bigcap_n\bigcup_\nu\{|f_{n+\nu}-f|\geqslant\varepsilon\}\right)=0.$$

特别当 μ 是有限测度时, $f_n\xrightarrow{\text{a.e.}}f$ 当且仅当对于任一 $\varepsilon>0$,

$$\mu\left(\bigcup_\nu\{|f_{n+\nu}-f|\geqslant\varepsilon\}\right)\to 0,\quad(n\to\infty).$$

ii) $f_{n+\nu}-f_n\xrightarrow{\text{a.e.}}0$ 当且仅当对于任一 $\varepsilon>0$

$$\mu\left(\bigcap_n\bigcup_\nu\{|f_{n+\nu}-f_n|\geqslant\varepsilon\}\right)=0.$$

特别当 μ 是有限测度时, $f_{n+\nu}-f_n\xrightarrow{\text{a.e.}}0$ 当且仅当对于任一 $\varepsilon>0$,

$$\mu\left(\bigcup_\nu\{|f_{n+\nu}-f_n|\geqslant\varepsilon\}\right)\to 0,(n\to\infty).$$

证 我们只证 i). ii) 的证明完全类似.

当 $f_n \xrightarrow{\text{a.e.}} f$ 时

$$\mu\left(\bigcap_n \bigcup_\nu |f_{n+\nu} - f| \geqslant \varepsilon\right) \leqslant \mu\left(\bigcup_{\varepsilon>0} \bigcap_n \bigcup_\nu \{|f_{n+\nu} - f| \geqslant \varepsilon\}\right)$$
$$= \mu(f_n \not\to f) = 0.$$

这就证明了条件的必要性

反之, 假定对任一 $\varepsilon > 0, \mu\left(\bigcap_n \bigcup_\nu \{|f_{n+\nu} - f| \geqslant \varepsilon\}\right) = 0$, 任取一收敛于零的正数序列 $\{\varepsilon_k\}$, 可得

$$\mu\{f_n \not\to f\} = \mu\left(\bigcup_k \bigcap_n \bigcup_\nu \{|f_{n+\nu} - f| \geqslant \varepsilon_k\}\right)$$
$$\leqslant \sum_k \mu\left(\bigcap_n \bigcup_\nu \{|f_{n+v} - f| \geqslant \varepsilon_k\}\right) = 0.$$

即 $f_n \xrightarrow{\text{a.e.}} f$. 这就证明了条件的充分性.

注意到 $A_n = \bigcup_\nu \{|f_{n+\nu} - f| \geqslant \varepsilon\}, n \geqslant 1$ 是一不增的集序列, 当 μ 是有限测度时, 利用测度的上连续性,

$$\mu\left(\bigcap_n A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

由此即得在 μ 是有限测度情况下 $f_n \xrightarrow{\text{a.e.}} f$ 的判别条件. □

从上面的证明中可以看出, 为了保证 $f_n \xrightarrow{\text{a.e.}} f$, 只要对某一收敛于零的正数序列 $\{\varepsilon_k\}, \mu\left(\bigcap_n \bigcup_\nu \{|f_{n+\nu} - f| \geqslant \varepsilon_k\}\right) = 0$ 就够了.

设 $(\Omega, \mathscr{A}, P)$ 是概率场, $\xi, \xi_n, n \geqslant 1$, 是这一概率场上的随机变量, 以上对于几乎处处收敛的讨论都适用于 $\xi, \{\xi_n\}$. 为了更切合概率的实际意义, 当 $\xi_n \xrightarrow{\text{a.e.}} \xi$ 时, 通常代替 "ξ_n 几乎处处收敛于 ξ" 而称为 "ξ_n 几乎必然收敛于 ξ". 记作 $\xi_n \xrightarrow{\text{a.s.}} \xi$. (当然此处的"几乎处处"和"几乎必然"都是关于 P 的.)

由于随机变量是有限值可测函数, 概率是有限测度, 所以随机变量序列 $\{\xi_n\}$ 几乎必然收敛于某一随机变量 ξ 当且仅当下列条件之一成立:

$$P\left(\bigcap_n \bigcup_\nu \{|\xi_{n+\nu} - \xi| \geqslant \varepsilon\}\right) = 0(\text{对任一}\varepsilon > 0);$$

$$P\left(\bigcup_\nu \{|\xi_{n+\nu} - \xi| \geqslant \varepsilon\}\right) \to 0(\text{对任一}\varepsilon > 0, \text{当}n \to \infty);$$

$$P\Big(\bigcap_n\bigcup_\nu\{|\xi_{n+\nu}-\xi_n|\geqslant\varepsilon\}\Big)=0(\text{对任一}\varepsilon>0);$$

$$P\Big(\bigcup_\nu\{|\xi_{n+\nu}-\xi_n|\geqslant\varepsilon\}\Big)\to 0(\text{对任一}\varepsilon>0,\text{当}n\to\infty);$$

2.5.2 依测度收敛

设 $(\Omega,\mathscr{A},\mu)$ 是任一给定的测度空间, 如无特别声明, 以下所说的可测函数都是这一测度空间上的可测函数.

定义 2 设 $\{f_n\}$ 是有限可测函数序列, f 是可测函数. 如果对任意 $\varepsilon>0$

$$\mu(\{|f_n-f|\geqslant\varepsilon\})\to 0,(n\to\infty),$$

则称序列 f_n **依测度 μ 收敛于** f, 记作

$$f_n\xrightarrow{\mu}f.$$

如果对于任意 $\varepsilon>0$,

$$\mu(\{|f_{n+\nu}-f_n|\geqslant\varepsilon\})\to 0,(\text{当}n\to\infty\text{时},\text{对}\nu\text{一致});$$

则称序列 f_n **依测度相互收敛**, 记作

$$f_{n+\nu}-f_n\xrightarrow{\mu}0.$$

由定义容易看出依测度收敛的极限函数是几乎处处有限的, 即若 $f_n\xrightarrow{\mu}f$, 则 fa.e. 有限. 事实上, 由定义 f_n 有限, 因此当 $f(\omega)=\pm\infty$ 时, $|f_n(\omega)-f(\omega)|=\infty$,

$$\begin{aligned}\mu(\{f=\pm\infty\})&=\mu(\{|f_n-f|=\infty\})\\&\leqslant\mu(\{|f_n-f|\geqslant\varepsilon\})\to 0.\end{aligned}$$

即 f, a.e. 有限.

测度收敛具有同普通收敛性相类似的性质.

性质 5 若 $f_n\xrightarrow{\mu}f,\{f_n'\}$ 为 $\{f_n\}$ 的任一子序列, 则 $f_n'\xrightarrow{\mu}f$.

性质 6 若 $f_n\xrightarrow{\mu}f,f_n\xrightarrow{\mu}f'$, 则 $f=f'$, a.e., 即依测度收敛的极限函数是几乎唯一的.

证 $\{f\neq f'\}\subset\{f=\pm\infty\}\cup\{f'=\pm\infty\}\cup\left(\bigcup\limits_{k=1}^{\infty}\left\{|f-f'|\geqslant\dfrac{1}{k}\right\}\right)$. 由于依测度收敛的极限函数是 a.e. 有限的, 因而

$$\mu(\{f=\pm\infty\})=\mu(\{f'=\pm\infty\})=0.$$

而由 $f_n \xrightarrow{\mu} f, f_n \xrightarrow{\mu} f'$ 知对任意 $\varepsilon > 0$,

$$\mu\{|f-f'| \geqslant \varepsilon\} \leqslant \mu\left\{|f-f_n| \geqslant \frac{\varepsilon}{2}\right\}$$

$$+\mu\left\{|f'-f_n| \geqslant \frac{\varepsilon}{2}\right\} \to 0,$$

即 $\mu(|f-f'| \geqslant \varepsilon) = 0$. 故得 $\mu(\{f \neq f'\}) = 0$, 即 $f = f'$,a.e. □

性质 7 若 $f_n \xrightarrow{\mu} f, g_n = f_n$(a.e.), $g = f$,(a.e), 则 $g_n \xrightarrow{\mu} g$.

证 由假设, 对任意 $\varepsilon > 0$,

$$\mu\{|f_n - f| \geqslant \varepsilon\} \to 0 (n \to \infty).$$

令 $N_n = \{g_n \neq f_n\}, N = \{g \neq f\}$, 则 $\mu(N_n) = \mu(N) = 0$, 取 $M = N \cup \left(\bigcup\limits_{n=1}^{\infty} N_n\right)$, 则 $\mu(M) = 0$, 在 M^c 上 $f_n = g_n, f = g$, 于是

$$\{|g_n - g| \geqslant \varepsilon\} = (\{|g_n - g| \geqslant \varepsilon\} \cap M^c)$$

$$\cup (\{|g_n - g| \geqslant \varepsilon\} \cap M) \subset \{|f_n - f| \geqslant \varepsilon\} \cup M.$$

因而

$$\mu(\{|g_n - g| \geqslant \varepsilon\}) \leqslant \mu(\{|f_n - f| \geqslant \varepsilon\}) + \mu(M) \to 0,$$

即 $g_n \xrightarrow{\mu} g$. □

由于依测度收敛的极限函数是 a.e 有限的, 再由性质 7, 我们可以假定依测度收敛的极限函数 (如果存在) 是有限的.

和 a.e. 收敛一样, 我们可以研究依测度收敛的运算性质, 为此, 我们证明下列定理及推论.

定理 3 设 $g(x_1, \cdots, x_m)$ 是定义在 $R^{(m)}$(或 $Z^{(m)}$) 中某一子集 D 上的一致连续函数, $f_k^{(n)}, f_k, k = 1, \cdots, m$ 是 $(\Omega, \mathscr{A})$ 上的可测函数, 对一切 $\omega \in \Omega, (f_1^{(n)}(\omega), \cdots, f_m^{(n)}(\omega)) \in D, (f_1(\omega), \cdots, f_m(\omega)) \in D$, 且 $f_k^{(n)} \xrightarrow{\mu} f_k, k = 1, \cdots, m$, 则

$$g(f_1^{(n)}, \cdots, f_m^{(n)}) \xrightarrow{\mu} g(f_1, \cdots, f_m).$$

证 只要证明任给 $\varepsilon > 0, \delta > 0$, 存在 $n_0(\varepsilon, \delta)$, 当 $n \geqslant n_0$ 时,

$$\mu(\{|g(f_1^{(n)}, \cdots, f_m^{(n)}) - g(f_1, \cdots, f_m)| \geqslant \varepsilon\}) < \delta.$$

由于 $g(x_1, \cdots, x_m)$ 在 D 上一致连续, 故存在 $\varepsilon_1(\varepsilon)$, 当 $(x_1', \cdots, x_m') \in D, (x_1, \cdots, x_m) \in D$ 且 $|x_k' - x_k| < \varepsilon_1, k = 1, \cdots, m$ 时

$$|g(x_1', \cdots, x_m') - g(x_1, \cdots, x_m)| < \varepsilon,$$

由于 $f_k^{(n)} \xrightarrow{\mu} f_k, k=1,\cdots,m$, 故对 $\varepsilon_1(\varepsilon)>0, \dfrac{\delta}{m}>0$, 存在 $n_k(\varepsilon,\delta)>0$, 当 $n>n_k$ 时,

$$\mu(|f_k^{(n)}-f_k|\geqslant\varepsilon_1)<\frac{\delta}{m}, k=1,\cdots,m.$$

令 $n_0=\max\limits_{1\leqslant k\leqslant m}\{n_k\}$, 当 $n>n_0$ 时

$$\begin{aligned}&\mu(\{|g(f_1^{(n)},\cdots,f_m^{(n)})-g(f_1,\cdots,f_m)|\geqslant\varepsilon\})\\ \leqslant&\mu\left(\bigcup_{k=1}^{m}\{|f_k^{(n)}-f_k|\geqslant\varepsilon_1\}\right)\\ \leqslant&\sum_{k=1}^{m}\mu(\{|f_k^{(n)}-f_k|\geqslant\varepsilon_1\})<\delta.\end{aligned}$$

于是有

$$g(f_1^{(n)},\cdots,f_m^{(n)})\xrightarrow{\mu}g(f_1,\cdots,f_m),\quad(n\to\infty).$$

推论 1 若 $f_n\xrightarrow{\mu}f, g_n\xrightarrow{\mu}g$, 则 $f_n\pm g_n\xrightarrow{\mu}f\pm g$.

证 设 $g(x_1,x_2)=x_1\pm x_2, D=R^{(2)}$(或 $Z^{(2)}$), g 在 D 上一致连续, 故此推论成立.

定理 4 设 μ 是 $(\Omega,\mathscr{A})$ 上的有限测度, $f_k^{(n)}, f_k, k=1,\cdots,m, n=1,2,\cdots$ 是其上可测函数且 $f_k^{(n)}\xrightarrow{\mu}f_k, k=1,\cdots,m$. 若 g 是 $R^{(m)}$(或 $Z^{(m)}$) 中开集 ① D 上的连续函数, 且对一切 $\omega\in\Omega$ 及 $n=1,2,\cdots, f^{(n)}(\omega)\triangleq(f_1^{(n)}(\omega),\cdots,f_m^{(n)}(\omega))\in D, f(\omega)\triangleq(f_1(\omega),\cdots,f_m(\omega))\in D$, 则

$$g(f_1^{(n)},\cdots,f_m^{(n)})\xrightarrow{\mu}g(f_1,\cdots,f_m),(n\to\infty).$$

证 取

$$D_N=\left\{x\in R^{(m)}:d(0,x)\leqslant N, d(x,D^c)\geqslant\frac{1}{N}\right\}.$$

其中 0 表示 $R^{(m)}$(或 $Z^{(m)}$) 中的原点 $(0,\cdots,0)$. $d(x,y)$ 表示 $R^{(m)}$(或 $Z^{(m)}$) 中的距离, $d(x,D^c)\triangleq\inf\limits_{y\in D^c}d(x,y)$. 而 $d(x,\varnothing)\triangleq\infty$. 显然 D_N 是非降的有界闭集序列; 且 $\bigcup\limits_{N=1}^{\infty}D_N=D$.

首先, 由于 $D\backslash D_N\downarrow\varnothing$, 故

$$f^{-1}(D_N^c)=\{\omega:f(\omega)\in D\backslash D_N\}\downarrow\varnothing.$$

① 关于开集的概念见本书第 6 章 6.3.1.

由 μ 是有限测度知, 任给 $\delta > 0$, 存在 $N(\delta)$, 使

$$\mu(f^{-1}(D_N^c)) < \frac{\delta}{2}.$$

其次, 由于 g 在有界闭集 D_{N+1} 上一致连续, 故对任意给定的 $\varepsilon > 0$, 存在 $\varepsilon_1(N(\delta),\varepsilon)$, 当 $|x'_k - x_k| < \varepsilon_1, k = 1,\cdots,m, x' = (x'_1,\cdots,x'_m) \in D_{N+1}, x = (x_1,\cdots,x_m) \in D_{N+1}$ 时有

$$|g(x') - g(x)| < \varepsilon.$$

任给 $\delta_1 > 0$, 存在 $x \in D_N, y \in D_{N+1}^c$ 使

$$d(D_N, D_{N+1}^c) > d(x,y) - \delta_1.$$

若 $d(0,y) > N+1$, 则 $d(x,y) \geqslant d(y,0) - d(x,0) \geqslant 1$; 若 $d(0,y) \leqslant N+1$, 则 $D^c \neq \varnothing, d(y,D^c) < \dfrac{1}{N+1}$, 于是存在 $z \in D^c$ 使 $d(y,z) < \dfrac{1}{N+1}$, 从而 $d(x,y) \geqslant d(x,z) - d(z,y) \geqslant \dfrac{1}{N} - \dfrac{1}{N+1} = \dfrac{1}{N(N+1)}$. 再由 δ_1 的任意性得

$$d(D_N, D_{N+1}^c) \geqslant \frac{1}{N(N+1)}.$$

取

$$\varepsilon_0 = \min\left\{\varepsilon_1, \frac{1}{\sqrt{m}N(N+1)}\right\},$$

就有

$$\{|g(f^{(n)}) - g(f)| \geqslant \varepsilon\} \subset \left(\bigcup_{k=1}^m \{|f_k^{(n)} - f_k| \geqslant \varepsilon_0\}\right) \cup f^{-1}(D_N^c).$$

事实上, 若

$$\omega \notin \bigcup_{k=1}^m \{|f_k^{(n)} - f_k| \geqslant \varepsilon_0\} \cup f^{-1}(D_N^c),$$

则

$$\omega \in \bigcap_{k=1}^m \{|f_k^{(n)} - f_k| < \varepsilon_0\} \cap f^{-1}(D_N),$$

则

$$f(\omega) \in D_N \subset D_{N+1} 且 \sum_{k=1}^m |f_k^{(n)}(\omega) - f_k(\omega)|^2 < m\varepsilon_0^2$$
$$\leqslant d^2(D_N, D_{N+1}^c),$$

则

$$f^{(n)}(\omega) \in D_{N+1} 且 |f_k^{(n)}(\omega) - f_k(\omega)| < \varepsilon_0 \leqslant \varepsilon_1,$$

$$k=1,\cdots,m,$$

则

$$|g(f^{(n)}(\omega))-g(f(\omega))|<\varepsilon,$$

则

$$\omega\notin\{|g(f^{(n)})-g(f)|\geqslant\varepsilon\}.$$

于是, 选 $n_0(\varepsilon_0,\delta)$, 使对一切 $k=1,\cdots,m$, 当 $n\geqslant n_0$ 时,

$$\mu(|f_k^{(n)}.-f_k|\geqslant\varepsilon_0)<\frac{\delta}{2m},$$

从而使

$$\mu(|g(f_1^{(n)},\cdots,f_m^{(n)})-g(f_1,\cdots,f_m)|\geqslant\varepsilon)$$
$$\leqslant\sum_{k=1}^{m}\mu(|f_k^{(n)}-f_k|\geqslant\varepsilon_0)+\mu(f^{-1}(D_N^c))<\delta.$$

此即

$$g(f_1^{(n)},\cdots,f_m^{(n)})\xrightarrow{\mu}g(f_1,\cdots,f_m),\quad(n\to\infty).$$ □

推论 2　若 μ 为有限测度, $f_k^{(n)}\xrightarrow{\mu}f_k,(n\to\infty),k=1,\cdots,m,R(x_1,\cdots,x_m)=\dfrac{P(x_1,\cdots,x_m)}{Q(x_1,\cdots,x_m)},P,Q$ 是多项式, 且对一切 $n\geqslant1$ 及 $\omega\in\Omega,Q(f_1^{(n)}(\omega),\cdots,f_m^{(n)}(\omega))\neq0,Q(f_1,(\omega),\cdots,f_m(\omega))\neq0$, 则 $R(f_1^{(n)},\cdots,f_m^{(n)})\xrightarrow{\mu}R(f_1,\cdots,f_m)$.

证　令

$$D=\{x\in R^{(m)}:Q(x)\neq0\},$$

(当 $f_k^{(n)},f_k$ 为复可测函数时以 $Z^{(m)}$ 代替 $R^{(m)}$), 则 D 是 $R^{(m)}(Z^{(m)})$ 中的开集. 而 $R(x)$ 是 D 上的连续函数, 从而由定理 4 知结论成立. □

下面的定理是说明几乎处处收敛与测度收敛之间的联系.

定理 5　设 $\{f_n\}$ 是有限可测函数序列.

i) 如果 $f_n\xrightarrow{\mu}f$, 则存在 $\{f_n\}$ 的子序列 $\{f_{n_k}\}$ 使 $f_{n_k}\xrightarrow{\text{a.e.}}f$.

ii) 如果 $f_{n+\nu}-f_n\xrightarrow{\mu}0$, 则存在 $\{f_n\}$ 的一个子序列 $\{f_{n_k}\},f_{n_k}\xrightarrow{\text{a.e.}}$ (某一)f, 且 $f_{nk}\xrightarrow{\mu}f$.

iii) 当 μ 是有限测度, 如果 $f_n\xrightarrow{\text{a.e.}}f$, 则 $f_n\xrightarrow{\mu}f$.

证　i) 因为 $f_n\xrightarrow{\mu}f$, 对于任一 $\varepsilon>0$,

$$\mu(|f_n-f|\geqslant\varepsilon)\to0,$$

所以对于每一正整数 k, 存在一个 n_k, 使得当 $n\geqslant n_k$ 时,

$$\mu\left(|f_n-f|\geqslant\frac{1}{2^k}\right)<\frac{1}{2^k},$$

可以假设 $n_1 < n_2 < \cdots < n_k < \cdots$，于是 $f'_k \triangleq f_{nk}, k = 1, 2, \cdots$ 是 $\{f_n\}$ 的一个子序列, 根据它的取法

$$\mu\left(|f'_k - f| \geqslant \frac{1}{2^k}\right) < \frac{1}{2^k},$$

这个子序列就几乎处处收敛于 f. 事实上, 对于任意的 $\varepsilon > 0$,

$$\begin{aligned}
&\mu\left(\bigcap_k \bigcup_\nu \{|f'_{k+\nu} - f| \geqslant \varepsilon\}\right)\\
\leqslant&\mu\left(\bigcup_\nu \{|f'_{k_0+\nu} - f| \geqslant \varepsilon\}\right)\text{对于每一个}k_0,\\
\leqslant&\sum_\nu \mu(\{|f'_{k_0+\nu} - f| \geqslant \varepsilon\})\\
\leqslant&\sum_\nu \mu\left(\left\{|f'_{k_0+\nu} - f| \geqslant \frac{1}{2^{k_0+\nu}}\right\}\right)\left(\text{对充分大的}k_0, \frac{1}{2^{k_0}} < \varepsilon\right)\\
\leqslant&\sum_\nu \frac{1}{2^{k_0+\nu}} = \frac{1}{2^{k_0}}.
\end{aligned}$$

由于上式对一切充分大的 k_0 都成立, 所以

$$\mu\left(\bigcap_k \bigcup_\nu \{|f'_{k+\nu} - f| \geqslant \varepsilon\right) = 0.$$

依据几乎处处收敛的判别条件即知

$$f'_k \xrightarrow{\text{a.e.}} f(\text{即} f_{n_k} \xrightarrow{\text{a.e.}} f).$$

ii) 由于 $f_{n+\nu} - f_n \xrightarrow{\mu} 0$, 故对于任意 $\varepsilon > 0$,

$\mu(\{|f_{n+\nu} - f_n| \geqslant \varepsilon\}) \to 0, (n \to \infty$ 时对 ν 一致成立).

因此, 对于每一正整数 k, 存在一个 n_k, 使得

$$\mu\left(\left\{|f_{n+\nu} - f_n| \leqslant \frac{1}{2^k}\right\}\right) < \frac{1}{2^k}, n \geqslant n_k, \nu \geqslant 1,$$

可以假定 $n_1 < n_2, < \cdots < n_k < \cdots$，若令 $f'_k = f_{n_k} (k = 1, 2, \cdots)$, 则 $\{f'_k\}$ 是 $\{f_n\}$ 的子序列, 由于 $\{f'_k\}$ 的取法, 有

$$\mu\left(\left\{|f'_{k+\nu} - f'_k| \geqslant \frac{1}{2^k}\right\}\right) < \frac{1}{2^k}, k \geqslant 1, \nu \geqslant 1.$$

我们证明这一子序列几乎处处收敛于某一 f, 为此令

$$A_k = \left\{|f'_{k+1} - f'_k| \geqslant \frac{1}{2^k}\right\}, \quad B_k = \bigcup_{l \geqslant k} A_l.$$

显然,

$$\mu(A_k) < \frac{1}{2^k}, \quad \mu(B_k) \leqslant \sum_{l \geqslant k} \mu(A_l) \leqslant \frac{1}{2^{k-1}},$$

对于任一给定的 $\varepsilon > 0$, 当 k 充分大时, 可使 $\frac{1}{2^{k-1}} < \varepsilon$, 这时

$$\{|f'_{k+\nu} - f'_k| \geqslant \varepsilon\} \subset B_k, \quad \nu = 1, 2, \cdots,$$

因而

$$\bigcup_{\nu} \{|f'_{k+\nu} - f'_k| \geqslant \varepsilon\} \subset B_k, \text{当}k\text{充分大}\left(\frac{1}{2^{k-1}} < \varepsilon\right).$$

包含关系 $\bigcup\limits_{\nu} \{|f'_{k+\nu} - f'_k| \geqslant \varepsilon\} \subset B_k$, 当 k 充分大时成立, 是因为: 若 $\omega \notin B_k$, 则 $\omega \notin A_l (l \geqslant k)$. 于是

$$\begin{aligned}|f'_{k+\nu}(\omega) - f'_k(\omega)| &\leqslant |f'_{k+1}(\omega) - f'_k(\omega)| + |f'_{k+2}(\omega) \\ &\quad - f'_{k+1}(\omega)| + \cdots + |f'_{k+\nu}(\omega) - f'_{k+\nu-1}(\omega)| \\ &\leqslant \frac{1}{2^k} + \frac{1}{2^{k+1}} + \cdots \leqslant \frac{1}{2^{k-1}} < \varepsilon.\end{aligned}$$

即 $\omega \notin \{|f'_{k+\nu} - f'_k| \geqslant \varepsilon\}$. 由此知

$$\begin{aligned}\mu\left(\bigcap_k \bigcup_{\nu} \{|f'_{k+\nu} - f'_k| \geqslant \varepsilon\}\right) &\leqslant \mu\left(\bigcup_{\nu} \{|f'_{k+\nu} - f'_k| \geqslant \varepsilon\}\right) \\ &\leqslant \mu(B_k) \leqslant \frac{1}{2^{k-1}}\end{aligned}$$

对一切充分大的 k 成立. 因而

$$\mu\left(\bigcap_k \bigcup_{\nu} \{|f'_{k+\nu} - f'_k| \geqslant \varepsilon\}\right) = 0, \text{对任一}\varepsilon > 0\text{成立}.$$

故 $f'_{k+\nu} - f'_k \xrightarrow{\text{a.e.}} 0$, 根据定理 1, 存在有限可测函数 f, 使 $f'_k \xrightarrow{\text{a.e.}} f$(即 $f_{n_k} \xrightarrow{\text{a.e.}} f$).

现在证明 $f'_k \xrightarrow{\mu} f$.

已知 $f'_k \xrightarrow{\text{a.e.}} f$, 存在 $N \in \mathscr{A}, \mu(N) = 0$, 使得当 $\omega \in N^c$ 时, $f'_k(\omega) \to f(\omega)(k \to \infty)$. 同时, 对于任一 $\varepsilon > 0$, 当 $\omega \in B_k^c$, $\frac{1}{2^{k-1}} < \varepsilon$ 时,

$$|f'_{k+\nu}(\omega) - f'_k(\omega)| < \varepsilon, \quad \nu \geqslant 1,$$

所以当 $\omega \in B_k^c \cap N^c$ 时, 由上面的两个关系式, 令 $\nu \to \infty$, 得

$$|f(\omega) - f'_k(\omega)| < \varepsilon,$$

即

$$\{|f-f'_k|<\varepsilon\}\supset B_k^c\cap N^c,\quad (k\text{充分大})$$

即

$$\{|f-f'_k|\geqslant\varepsilon\}\subset B_k\cup N.$$

因此

$$\mu(\{|f-f'_k|\geqslant\varepsilon\})\leqslant\mu(B_k)+\mu(N)<\frac{1}{2^{k-1}}\to 0,\quad (k\to\infty).$$

即 $f'_n\xrightarrow{\mu}f$

iii) 假定 $f_n\xrightarrow{\text{a.e.}}f$, 当 μ 有限时, 利用定理 2 i).

$$\mu(\{|f_{n+1}-f|\geqslant\varepsilon\})\leqslant\mu\left(\bigcup_{\nu}\{|f_{n+\nu}-f|\geqslant\varepsilon\}\right)\to 0,$$

故知 $f_n\xrightarrow{\mu}f$.

最后, 我们给出测度收敛的判定准则.

定理 6 $f_n\xrightarrow{\mu}$(某一)f 的充分必要条件是 $f_{n+\nu}-f_n\xrightarrow{\mu}0$(对 ν 一致).

证 设 $f_n\xrightarrow{\mu}f$, 则由于对任意 $\varepsilon>0$,

$$\mu(\{|f_{n+\nu}-f_n|\geqslant\varepsilon\})\leqslant\mu\left(\left\{|f_{n+\nu}-f|\geqslant\frac{\varepsilon}{2}\right\}\right)+\mu\left(\left\{|f_n-f|\geqslant\frac{\varepsilon}{2}\right\}\right),$$

易见

$$\mu(\{|f_{n+\nu}-f_n|\geqslant\varepsilon\})\to 0(\text{当}n\to\infty,\text{对}\nu\text{一致}),$$

即

$$f_{n+\nu}-f_n\xrightarrow{\mu}0,\qquad(\text{对}\nu\text{一致})$$

反之, 如果 $f_{n+\nu}-f_n\xrightarrow{\mu}0(n\to\infty$, 对 ν 一致), 则由定理 5ii), 存在一个子序列 $\{f_{n_k}\}$ 及一有限可测函数 f, 使得 $f_{nk}\xrightarrow{\mu}f$, 则

$$\mu(\{|f-f_n|\geqslant\varepsilon\})\leqslant\mu\left(\left\{|f-f_{n_k}|\geqslant\frac{\varepsilon}{2}\right\}\right)+\mu\left(\left\{|f_{n_k}-f_n|\geqslant\frac{\varepsilon}{2}\right\}\right),$$

在上式中取 $n_k>n$, 令 $n\to\infty$, 即得

$$\mu(\{|f-f_n|\geqslant\varepsilon\})\to 0,$$

即 $f_n \xrightarrow{\mu} f(n \to \infty)$.

特别, 当测度 μ 是概率时, 依测度收敛的概念就成为**依概率收敛**的概念. 本节的结论对于依概率收敛自然都适用. 此外, 如在定义 2 中将 $\{f_n\}$ 有限改为 a.e. 有限, 一切结论仍然成立.

2.5.3 分布律收敛

定义 3 设随机变量 ξ_n 的分布函数是 $F_n, n = 1, 2, \cdots$, 随机变量 ξ 的分布函数是 F, 若对 F 的每一连续点 x_0 来说, $F_n(x_0) \to F(x_0)$, 则称 ξ_n 的分布律收敛于 ξ 的分布律, 记作 $F_n \xrightarrow{C} F$ 或 $\mathscr{L}(\xi_n) \to \mathscr{L}(\xi)$.

注意, 在现在的情况下, 自然有 $F_n(+\infty) \to F(+\infty), F_n(-\infty) \to F(-\infty)$.

现在我们讨论依分布律收敛性及与其他收敛性之间的关系.

定理 7 如果 $\xi_n \xrightarrow{P} \xi$, 则 $\mathscr{L}(\xi_n) \to \mathscr{L}(\xi)$, 即

$$F_n \xrightarrow{C} F.$$

证 需要证明对于 F 的每一连续点 x,

$$F_n(x) \to F(x).$$

首先, 对于任意 $x' > x$,

$$\begin{aligned}\{\xi_n < x\} &= \{\xi_n < x, \xi < x'\} + \{\xi_n < x, \xi \geqslant x'\} \\ &\subset \{\xi < x'\} \cup \{|\xi_n - \xi| \geqslant x' - x\},\end{aligned}$$

取概率得

$$F_n(x) \leqslant F(x') + P(\{|\xi_n - \xi| \geqslant x' - x\}),$$

由于 $\xi_n \xrightarrow{P} \xi, x' - x > 0$, 故 $P(\{|\xi_n - \xi| \geqslant x' - x\}) \to 0$, 因而

$$\overline{\lim_{n\to\infty}} F_n(x) \leqslant F(x').$$

再令 $x' \to x$, 注意到 x 是 F 的连续点, 得出

$$\overline{\lim_{n\to\infty}} F_n(x) \leqslant F(x). \tag{3}$$

类似地, 对于任意 $x'' < x$, 有

$$\begin{aligned}&\{\xi < x''\} \subset \{\xi_n < x\} \cup \{|\xi_n - \xi| \geqslant x - x''\}, \\ &F(x'') \leqslant F_n(x) + P(\{|\xi_n - \xi| \geqslant x - x''\}), \\ &F(x'') \leqslant \underline{\lim_{n\to\infty}} F_n(x),\end{aligned}$$

$$F(x) \leqslant \varliminf_{n\to\infty} F_n(x). \tag{4}$$

由 (3) 和 (4) 知, 对于 F 的每一连续点 x,

$$\lim_{n\to\infty} F_n(x) = F(x).$$

□

推论 3 $\xi_n \xrightarrow{P} a$(常数), 当且仅当 $\mathscr{L}(\xi_n) \to \mathscr{L}(a)$, 即 $F_{\xi_n} \xrightarrow{C} F$, 其中

$$F(x) = \begin{cases} 0, & x \leqslant a, \\ 1, & x > a. \end{cases}$$

证 根据定理 7, 当 $\xi_n \xrightarrow{P} a$ 时, 显然有 $\mathscr{L}(\xi_n) \to \mathscr{L}(a)$, 反之, 假定 $\mathscr{L}(\xi_n) \to \mathscr{L}(a)$, 则对任一 $\varepsilon > 0$,

$$\begin{aligned} P(\{|\xi_n - a| \geqslant \varepsilon\}) &= P(\{\xi_n \leqslant a - \varepsilon\}) + P(\{\xi_n \geqslant a + \varepsilon\}) \\ &\leqslant P\left(\left\{\xi_n < a - \frac{\varepsilon}{2}\right\}\right) + 1 - P(\{\xi_n < a + \varepsilon\}). \\ &= F_{\xi_n}\left(a - \frac{\varepsilon}{2}\right) + 1 - F_{\xi_n}(a + \varepsilon) \to 0. \end{aligned}$$

□

用定理 7 的论证方法, 可以得到

定理 8 如果 $\xi_n - \xi_n' \xrightarrow{P} 0$ 且 $\mathscr{L}(\xi_n') \to \mathscr{L}(\xi)$, 则 $\mathscr{L}(\xi_n) \to \mathscr{L}(\xi)$.

定理 9(Слуцкий). 如果 $\mathscr{L}(\xi_n) \to \mathscr{L}(\xi), \mathscr{L}(\eta_n) \to \mathscr{L}(a)$, 则 $\mathscr{L}(\xi_n + \eta_n) \to \mathscr{L}(\xi + a)$.

证 由推论 3 知 $\mathscr{L}(\eta_n) \to \mathscr{L}(a)$ 等价于 $\eta_n \xrightarrow{P} a$. 再由

$$(\xi_n + \eta_n) - (\xi_n + a) \xrightarrow{P} 0$$

及 $\mathscr{L}(\xi_n) \to \mathscr{L}(\xi)$, 经过简单的推理可得

$$\mathscr{L}(\xi_n + a) \to \mathscr{L}(\xi + a).$$

于是由定理 8 即知

$$\mathscr{L}(\xi_n + \eta_n) \to \mathscr{L}(\xi + a).$$

关于分布律收敛的进一步结果将要在第 6 章讨论.

习题及补充

1. 若测度 μ 有限, 则对给定的有限可测函数 f 和任一 $\varepsilon > 0$, 存在集合 A, 使 $\mu(A) < \varepsilon$, 并且 f 在 A^c 上有界.

定义　称 f_n 在集 A 上一致收敛向 f, 如果对一切 $\omega \in A$, 存在 n_0, 当 $n \geqslant n_0$ 时

$$\begin{aligned} &|f_n(\omega) - f(\omega)| < \varepsilon, && 若|f(\omega)| < \infty, \\ &f_n(\omega) < -\frac{1}{\varepsilon}, && 若f(\omega) = -\infty, \\ &f_n(\omega) > \frac{1}{\varepsilon} && 若f(\omega) = +\infty. \end{aligned}$$

2. 若在 Ω 上 $f_n \xrightarrow{一致} f$, 则 $f_n \xrightarrow{\text{a.e.}} f$, 但未必有 $f_n \xrightarrow{\mu} f$. 试举例说明 (不假定 fa.e. 有限).

定义　如果对任一 $\varepsilon > 0$, 存在集 $A, \mu(A) < \varepsilon$, 并且在 A^c 上 $f_n \xrightarrow{一致} f$, 便称序列 f_n 几乎一致收敛于 f, 记作 $f_n \xrightarrow{\text{a.u.}} f$.

3. 设 f_n, fa.e. 有限, 试证:

i) 若 $f_n \xrightarrow{\text{a.u.}} f$ 则 $f_n \xrightarrow{\text{a.e.}} f$;

ii) 若 $f_n \xrightarrow{\mu} f$, 则存在子序列 $f_n \xrightarrow{\text{a.u.}} f$.

4. 若 μ 有限, 则 $f_n \xrightarrow{\text{a.e.}} f$ 蕴含 $f_n \xrightarrow{\text{a.u.}} f$.

5. 若 μ 是 $\sigma-$ 有限测度, f_n, f 有限, $f_n \xrightarrow{\text{a.e.}} f$, 则存在 $N, \mu(N) = 0$, 存在 $A_j, j = 1, 2, \cdots$ 两两不交, 使 $N^c = \sum_{j=1}^{\infty} A_j$, 且在每一 A_j 上 $f_n \xrightarrow{\text{a.u.}} f$.($N, A_j, j = 1, 2, \cdots$ 皆为测度空间上的可测集.)

6. 若 μ 有限, $f_n \xrightarrow{\text{a.e.}} f, f, f_n$ 是有限可测函数, 则对任意给定的 $\varepsilon > 0$, 存在集 $A, \mu(A) > \mu(\Omega) - \varepsilon$, 且在 A 上 f_n 一致有界.

7. 若 μ 有限, $f_{mn} \xrightarrow{\text{a.e.}} f_m (n \to \infty), f_m \xrightarrow{\text{a.e.}} f (m \to \infty)$, 则有子序列 $\{m_k\}, \{n_k\}$ 使 $f_{m_k n_k} \xrightarrow{\text{a.e.}} f, (k \to \infty)$ (假定 f_{mn}, f_m, f 均 a.e. 有限).

8. 不用定理 3, 直接证明定理 3 的推论 1.

9. 直接证明定理 4 的推论 2.

10. 证明定理 8.

11. 若 $\mathscr{L}(\xi_n) \to \mathscr{L}(\xi)$, 试证 $\mathscr{L}(\xi_n + a) \to \mathscr{L}(\xi + a)$, 其中 a 是任一有限常数.

12. 若 $\mathscr{L}(\xi_n) \to \mathscr{L}(\xi), \mathscr{L}(\eta_n) \to \mathscr{L}(a)$ 且对一切 $n, P(\eta_n = 0) = 0$ 试证

i) $\mathscr{L}(\xi_n \eta_n) \to \mathscr{L}(a\xi) (a > 0)$;

ii) $\mathscr{L}(\xi_n / \eta_n) \to \mathscr{L}\left(\dfrac{\xi}{a}\right) (a > 0)$.

第 3 章　数学期望与积分

§3.1 引　　言

我们知道, 随机变量的分布函数已经完全地描述了随机变量的概率规律. 但某些数字特征如数学期望、方差、相关矩及高阶矩等则更集中地反映了随机变量的特征, 要对它们作理论探讨, 只利用普通数学分析的工具是不够的, 需要用到一般测度空间上的积分理论. 因此, 在本章中我们将介绍这种积分理论, 并用来研究数学期望、方差和矩的性质. 在本章中我们也将利用积分的理论对于分布函数作进一步的研究.

现在, 我们从数学期望的实际意义出发.

例 1　假设用一种武器向同一目标进行射击, 用一颗子弹即可击中目标的概率为 p_1, 用两颗子弹击中目标的概率为 $p_2, \cdots$ 此外知道至多用 n 颗子弹一定能打中目标, 即 $p_1 + \cdots + p_n = 1$. 试问击中这一目标平均需要多少颗子弹?

设需要 k 粒子弹击中目标的事件为 A_k, 则 $P(A_k) = p_k, k = 1, \cdots, n$. 用 ξ 表示击中一次目标所需的子弹数, 则

$$\xi = 1\chi_{A_1} + \cdots + n\chi_{A_n}$$

是一个简单随机变量. 问题是求 ξ 的平均取值是多少. 根据概率的实际意义, 容易知道击中这一目标平均需要的子弹数是

$$1p_1 + \cdots + np_n.$$

一般地, 对任一简单随机变量, 它所取的可能值为 $x_1, \cdots, x_n$, 取这些值相应的概率为 $p_1, \cdots, p_n$, 我们称

$$\begin{aligned} E\xi &\widehat{=} x_1 p_1 + \cdots + x_n p_n \\ &= x_1 P(\xi = x_1) + \cdots + x_n P(\xi = x_n) \end{aligned}$$

为 ξ 的数学期望, 它表示 ξ 取值的平均状况, 所以有时也称它为 ξ 的平均值. 不过, 它与 n 个数的算术平均值不同, 它相当于加权平均.

例 2　设 ξ 是在时间间隔 $(a, a+t]$ 内某电话局收到用户呼唤的次数, ξ 所取的一切可能值为 $0, 1, \cdots, n, \cdots$, 且 $P(\xi = n) = e^{-\lambda t}\dfrac{(\lambda t)^n}{n!}$, 为了估计呼唤的频繁情况, 就需要求出 ξ 的平均值.

此例中的 ξ 为初等随机变量, 它是简单随机变量的极限. 设收到 n 次呼唤的事件为 A_n, 则

$$\xi=\sum_{n=1}^{\infty}n\chi_{A_n}.$$

若令

$$\xi_n=\sum_{k=1}^{n}k\chi_{A_k},\qquad P(\xi_n=0)=1-\sum_{k=1}^{n}P(\xi=k),$$

于是 $\lim\limits_{n\to\infty}\xi_n=\xi$. 自然我们认为 ξ 的平均值是 ξ_n 的平均值的极限, 即

$$\begin{aligned}E\xi&=\lim_{n\to\infty}E\xi_n=\lim_{n\to\infty}\sum_{k=1}^{n}ke^{-\lambda t}\frac{(\lambda t)^k}{k!}\\&=\sum_{k=1}^{\infty}ke^{-\lambda t}\frac{(\lambda t)^k}{k!}=\lambda t.\end{aligned}$$

于是我们得知 λ 表示在单位时间内某电话局平均收到的呼唤次数.

一般地, 设 ξ 是一个离散型随机变量, 它所取的一切可能值为 $x_1,\cdots x_n,\cdots$, 就称

$$E\xi\triangleq\sum_{n=1}^{\infty}x_nP(\xi=x_n)$$

为 ξ 的数学期望.(以后我们要指出 $E\xi$ 存在的条件.)

现在我们看对一般的随机变量 ξ 来说, 它的平均值应该怎样来描述.(其严格定义见本章 §4.)

我们知道随机变量 ξ 是给定概率场上的有限可测函数, 它可以表成这个概率场上某一简单函数序列 $\{\xi_n\}$ 的极限. 我们认为 ξ 的数学期望应当是

$$E\xi=\lim_{n\to\infty}E\xi_n.$$

为了看清楚这样的 $E\xi$ 的确表示 ξ 的平均值, 我们不妨取出一个具体的简单函数序列 $\{\xi_n\}$, 并把上面的极限等式明显地写出来.

以 ξ 是非负随机变量为例. 取

$$\xi_n(\omega)=\begin{cases}\dfrac{k}{2^n},\omega\in\left\{\dfrac{k}{2^n}\leqslant\xi(\omega)<\dfrac{k+1}{2^n}\right\},k=0,1,\cdots,n2^n-1,\\ n,\omega\in\{\xi(\omega)\geqslant n\}.\end{cases}$$

显然 ξ_n 是简单函数, 且 $\lim\limits_{n\to\infty}\xi_n=\xi$.

$$E\xi_n=\sum_{k=0}^{n2^n-1}\frac{k}{2^n}P\left(\frac{k}{2^n}\leqslant\xi<\frac{k+1}{2^n}\right)+nP(\xi\geqslant n).$$

从这里我们看到很类似于黎曼积分求面积和体积的方法. 我们把 ξ 所取的值分成若干个小区间, 将 ξ 取在同一区间的值算作是相同的, 例如当 $\frac{k}{2^n} \leqslant \xi < \frac{k+1}{2^n}$ 时, 算作 $\xi = \frac{k}{2^n}, k = 0, 1, \cdots, n2^n - 1$,而当 $\xi \geqslant n$ 时, 算作 $\xi = n$. 于是我们就可以得到 ξ 的平均值的一个近似值, 它正好就是 $E\xi_n$, 而且随着 n 的增大它愈来愈精确, 所以我们定义 $E\xi = \lim\limits_{n\to\infty} E\xi_n$ 是有道理的.

这里的 $\frac{k}{2^n}$ 类似于黎曼积分和 $\Sigma f(x_k')\Delta x_k$ 中的 $f(x_k')$, 而 $P\left(\frac{k}{2^n} \leqslant \xi < \frac{k+1}{2^n}\right)$ 相当于 Δx_k. 但是它们之间还是有原则的区别, 即在黎曼积分中是将函数的定义域划分成小区间, 这里我们则是将函数值域划分成小区间, 前者只有在充分小的区间上函数值的改变非常小时, 才可能有极限, 因此黎曼积分基本上只适用于连续函数, 而现在我们所涉及的函数是随机变量, 它的定义域是基本事件集 Ω, 因此按定义域划分的原则就很难采用. 首先将 Ω 划分为可测集的和, 而让这些集的测度趋向于零, 这种划分方法未必存在; 其次, 即使这种划分是可能的, 如上所述, 也仍然带有局限性, 因为这样也只适用于那些在测度充分小的集上 ξ 的值改变很小的情况. 看来, 要建立数学期望的概念并讨论它的性质, 只利用普通数学分析的工具就显得不够了. 我们将会看到在抽象空间中的积分采用将函数值域划分成小区间的方法有着广阔的适用范围. 这样一来, 随机变量 ξ 的数学期望的概念将等同于在概率空间 $(\Omega, \mathscr{A}, P)$ 上 ξ 的积分的概念.

§3.2 积分的定义和性质

§1 我们指出随机变量 ξ 的数学期望应是概率空间 $(\Omega, \mathscr{A}, P)$ 上随机变量 ξ 的积分. 这一节中, 我们将给出一般测度空间上可测函数的积分的定义, 并研究它的基本性质.

如无相反的说明, 下面出现的集 $A, B, \cdots$, 函数 $f, g, \cdots$ 分别为取定的 (但是任意的) 测度空间 $(\Omega, \mathscr{A}, \mu)$ 上的可测集和可测函数.

3.2.1 积分的定义

我们首先定义非负简单函数的积分, 然后定义非负可测函数的积分, 再次定义一般实可测函数的积分, 最后定义复可测函数的积分.

定义 1 设 f 是非负简单函数, $f = \sum\limits_{k=1}^{m} x_k \chi_{A_k}, x_k \geqslant 0, A_1, \cdots, A_m$ 两两不交, $\sum\limits_{k=1}^{m} A_k = \Omega$, 则称值 $\sum\limits_{k=1}^{m} x_k \mu(A_k)$ 为 f 在 Ω 上关于测度 μ 的积分 (简称 f 对 μ 的

积分或 f 的积分), 记作 $\int_\Omega f(\omega)d\mu(\omega)$, 简记作 $\int fd\mu$ 或 $\int f$(在不发生混淆的情况下). 若 $A\in\mathscr{A}$, 称 $\int f\chi_A d\mu$ 为 f 在集 A 上的积分, 记作 $\int_A fd\mu$.

显然按照这一定义, 任意非负简单函数的积分一定存在 (不过可能是 $+\infty$). 但我们还需证明它是唯一确定的, 与 f 的表示法无关. 即需证: 假如 f 有另一表示式 $f=\sum\limits_{j=1}^n y_j\chi_{B_j}, y_j\geqslant 0, B_j, j=1,\cdots,n$ 两两不交, $\sum\limits_{j=1}^n B_j=\Omega$, 则

$$\sum_{k=1}^m x_k\mu(A_k)=\sum_{j=1}^n y_j\mu(B_j).$$

事实上, 因为 $\sum\limits_{j=1}^n B_j=\Omega, \sum\limits_{k=1}^m A_k=\Omega$, 故有 $A_k=\sum\limits_{j=1}^n A_k\cap B_j, B_j=\sum\limits_{k=1}^m A_k\cap B_j$, 而且在 $A_k\cap B_j\neq\varnothing$ 时 $x_k=y_j$, 当 $A_k\cap B_j=\varnothing$ 时 $\mu(A_k\cap B_j)=0$. 因此

$$\begin{aligned}\sum_{k=1}^m x_k\mu(A_k)&=\sum_{k=1}^m x_k\sum_{j=1}^n\mu(A_k\cap B_j)\\&=\sum_{k=1}^m\sum_{j=1}^n y_j\mu(A_k\cap B_j)\\&=\sum_{j=1}^n\sum_{k=1}^m y_j\mu(A_k\cap B_j)\\&=\sum_{j=1}^n y_j\mu(B_j).\end{aligned}$$ □

为了合理地给出非负可测函数积分的定义, 先研究非负简单函数积分的几个性质.

性质 1　若 f,g 是非负简单函数, $f\leqslant g$, 则 $\int f\leqslant\int g$.

证　设

$$f=\sum_{k=1}^m x_k\chi_{A_k},\quad g=\sum_{k=1}^n y_j\chi_{B_j},$$

由于 $f\leqslant g$, 则在 $A_k\cap B_j\neq\varnothing$ 时 $x_k\leqslant y_j$, 而当 $A_k\cap B_j=\varnothing$ 时 $\mu(A_k\cap B_j)=0$.

$$\begin{aligned}\int fd\mu&=\sum_{k=1}^m x_k\mu(A_k)=\sum_{k=1}^m\sum_{j=1}^n x_k\mu(A_k\cap B_j)\\&\leqslant\sum_{j=1}^n\sum_{k=1}^m y_j\mu(A_k\cap B_j)=\sum_{j=1}^n y_j\sum_{k=1}^m\mu(A_k\cap B_j)\\&=\sum_{j=1}^n y_j\mu(B_j)=\int gd\mu.\end{aligned}$$ □

性质 2 若 $c \geqslant 0, f$ 是非负简单函数, 则

$$\int cf d\mu = c\int f d\mu.$$

(证明从略).

性质 3 若 f 是给定的非负简单函数, 令

$$\varphi(A) = \int_A f d\mu.$$

则 φ 为 $\mathscr{A}$ 上 σ-可加集函数.

证 设 $f = \sum_{k=1}^{m} x_k \chi_{A_k}$, 则

$$\begin{aligned}\varphi(A) &= \int_A f d\mu = \int (f\chi_A + 0\cdot\chi_{A^c})d\mu \\ &= \sum_{k=1}^{m} x_k \mu(A\cap A_k) + 0\cdot\mu(A^c) \\ &= \sum_{k=1}^{m} x_k \mu(A\cap A_k).\end{aligned}$$

由于 μ 是 σ-可加的, 因而 φ 也是 σ-可加的. □

定义 2 若 f 是非负可测函数, 则定义 f 在 Ω 上对 μ 的积分

$$\int f d\mu \widehat{=} \sup\left\{\int s d\mu : 0\leqslant s\leqslant f, s\text{为简单函数}\right\}. \tag{1}$$

对 $A\in\mathscr{A}$, 定义 f 在集 A 上的积分

$$\int_A f d\mu \widehat{=} \int f\chi_A d\mu$$

首先我们看到对一切非负可测函数, 由 (1) 式所定义的积分是存在的、确定的, 并且由性质 1 知非负简单函数的积分也符合 (1) 式. 因此定义 2 是定义 1 的拓广.

对于非负可测函数, 性质 1 仍然成立, 即

性质 1′ 若 $0\leqslant f\leqslant g$, 则 $\int f d\mu \leqslant \int g d\mu$.

证 对任意满足 $0\leqslant s\leqslant f$ 的简单函数 s, 必有 $0\leqslant s\leqslant g$, 因而

$$\left\{\int s d\mu : 0\leqslant s\leqslant f, s\text{简单函数}\right\} \subset \left\{\int s d\mu : 0\leqslant s\leqslant g, s\text{简单函数}\right\}.$$

故由定义 2 知

$$\int f d\mu \leqslant \int g d\mu. \qquad \square$$

下面我们叙述并证明在积分论中起着重要作用的定理.

定理 1(单调收敛定理)　若 $\{f_n\}$ 是非负可测函数序列, 且 $f_n \uparrow f$, 则

$$\lim_{n\to\infty}\int f_n d\mu = \int f d\mu.$$

证　由性质 $1'$, $\int f_n d\mu$ 是非降序列, 因而极限存在 (可能是 ∞), 记作 a. 又由于对一切 $n \geqslant 1$, $f_n \leqslant f$ 及性质 $1'$, 我们有

$$a \leqslant \int f d\mu.$$

另一方面, 令 $0 \leqslant s \leqslant f$, s 是简单函数, 设

$$s = \sum_{j=1}^{m} x_j \chi_{A_j} + \infty\chi_{A_{m+1}}, 0 \leqslant x_j < \infty, j = 1, \cdots, m,$$

$$s_{c,k} = \sum_{j=1}^{m} c x_j \chi_{A_j} + k\chi_{A_{m+1}},$$

其中 $0 < c < 1, k$ 是正整数. 又令

$$\Omega_n = \{\omega : s_{c,k}(\omega) \leqslant f_n(\omega)\},$$

容易看出 $\Omega_n \in \mathscr{A}, \Omega_n \uparrow \Omega$. 由于

$$f_n \geqslant f_n\chi_{\Omega_n} \geqslant s_{c,k}\chi_{\Omega_n},$$

利用性质 $1'$ 知

$$\int f_n d\mu \geqslant \int s_{c,k}\chi_{\Omega_n} d\mu = \int_{\Omega_n} s_{c,k} d\mu.$$

由性质 3 知 $\varphi(A) = \int_A s_{c,k} d\mu$ 是 σ-可加集函数, 因而是下连续的, 从而令 $n \to \infty$ 得到

$$a \geqslant \int_{\Omega} s_{c,k} d\mu = \sum_{j=1}^{m} c x_j \mu(A_j) + k\mu(A_{m+1}).$$

再令 $c \to 1, k \to \infty$, 得

$$a \geqslant \sum_{j=1}^{m} x_j \mu(A_j) + \infty\mu(A_{m+1}) = \int s d\mu.$$

在 $0 \leqslant s \leqslant f$ 的简单函数 s 的类上对上式右端取上确界, 即得

$$a \geqslant \int f d\mu. \qquad \square$$

根据定理 1, 对于非负可测函数 f, 任意两个满足 $0 \leqslant f_n \uparrow f, 0 \leqslant g_n \uparrow f$ 的可测函数 (特别可以取为简单函数) 序列 $\{f_n\}$, $\{g_n\}$, 都有

$$\lim_{n\to\infty} \int f_n d\mu = \lim_{n\to\infty} \int g_n d\mu.$$

而任意非负可测函数可以表成非降的非负简单函数序列的极限, 这样就使我们得出非负可测函数积分的另一等价定义.

定义 2′ 若 f 是非负可测函数, $\{f_n\}$ 是具有性质 $0 \leqslant f_n \uparrow f$ 的简单函数序列, 定义 f 在 Ω 上对 μ 的积分

$$\int f d\mu \widehat{=} \lim_{n\to\infty} \int f_n d\mu.$$

定义 f 在集 $A \in \mathscr{A}$ 上的积分

$$\int_A f d\mu = \int f \chi_A d\mu.$$

这一定义使用起来比定义 2 方便.

定义 3 设 f 是实可测函数, f^+ 和 f^- 分别是它的正部和负部, 若 $\int f^+ d\mu$, $\int f^- d\mu$ 至少有一有限, 则定义 f 在 Ω 上对 μ 的积分

$$\int f d\mu \widehat{=} \int f^+ d\mu - \int f^- d\mu.$$

并称 **f 的积分存在**. 定义 f 在集 A 上对 μ 的积分

$$\int_A f d\mu \widehat{=} \int f \chi_A d\mu.$$

如果 $\int f d\mu$ 有限, 则称 f 在 Ω 上**可积**.

定义 4 若 f 是复可测函数, $f = f_1 + i f_2, f_1, f_2$ 是实可测函数, 若 f_1, f_2 的积分存在, 则称

$$\int f d\mu \widehat{=} \int f_1 d\mu + i \int f_2 d\mu$$

为 f 在 Ω 上对 μ 的积分. 若积分有限, 称 f 在 Ω 上可积. 称

$$\int_A f d\mu = \int f \chi_A d\mu$$

为 f 在集 $A \in \mathscr{A}$ 上的积分.

为了对于 a.e. 可测的函数给出积分的定义, 我们先证明

性质 4 若 $f = g$, a.e., 则 $\int f d\mu = \int g d\mu$.

证 若 f, g 是非负简单函数, 设 $f = g$, a.e. 且

$$f = \sum_{k=1}^{m} x_k \chi_{A_k}, g = \sum_{j=1}^{n} y_1 \chi_B,$$

故在 $\mu(A_k \cap B_j) \neq 0$ 时 $x_k = y_j$, 从而 $\int f d\mu = \int g d\mu$.

若 f 和 g 是非负可测函数, 则对任何满足 $0 \leqslant s \leqslant f$ 的简单函数 s, 必存在简单函数 $s' = s$, a.e., 且 $0 \leqslant s' \leqslant g$, 从而 $\int s d\mu = \int s' d\mu$, 因此 $\int f d\mu \leqslant \int g d\mu$. 同理可证 $\int f d\mu \geqslant \int g d\mu$. 即 $\int f d\mu = \int g d\mu$.

若 f, g 为实函数, $f = g$, a.e., 必有 $f^{\pm} = g^{\pm}$, a.e.. 若 f, g 是复函数, $f = g$, a.e., 则必有 $f_i = g_i$, a.e., $i = 1, 2. f_i, g_i, i = 1, 2$ 分别为 f, g 的实、虚部. 故性质 4 对积分存在的任意实、复可测函数均成立.

定义 5 若 f a.e. 可测 (或 a.e. 有定义), $f = g$, a.e., 且 g 可测, 积分存在, 则定义 f 在 Ω 上对 μ 的积分为

$$\int f d\mu \widehat{=} \int g d\mu.$$

由于性质 4 及 a.e. 相等关系具有传递性, 可知定义 5 中 $\int f d\mu$ 是确定的.(不因 g 的不同选择而异.)

3.2.2 积分的性质

定理 2 可测函数的积分具有下列性质:

1) 线性性质:

a) 若 $\int f, \int g, \int f + \int g$ 存在, 则 $\int (f+g) = \int f + \int g$;

b) 若 $\int f$ 存在且 $A \cap B = \varnothing$, 则 $\int_{A+B} f = \int_A f + \int_B f$;

c) $\int cf = c \int f$, 若 f 是实函数, 则要求 c 有限, 若 f 是复函数, 还要求 f 可积;

d) $\int \overline{f} = \overline{\int f}$.

2) 单调性质:

a) f, g 是实函数, $\int f, \int g$ 存在, 若 $f \geqslant g$, a.e., 则对一切 $A \in \mathscr{A}$, $\int_A f \geqslant \int_A g$; 反之, 若 μ 是 σ-有限的, 则逆命题成立;

b) $\int f$ 存在, 则 $\left|\int f\right| \leqslant \int |f|$;

c) $f \geqslant 0$, 则 $\int f = 0$ 的充分必要条件是 $f = 0$, a.e..

3) 可积性质:

a) f 可积的充分必要条件是 $\int |f| < \infty$, 当 f 可积时, f a.e. 有限;

b) $|f| \leqslant g, g$ 可积, 则 f 可积;

c) f, g 可积, 则 $f + g$ 可积;

d) 若 $\int (fg)$ 存在, 则 $\left|\int (fg)\right|^2 \leqslant \int |f|^2 \cdot \int |g|^2$.

证 1) 线性性质:

a) 首先对 f, g 是非负简单函数的情形进行证明: 设

$$f = \sum_{k=1}^{m} x_k \chi_{A_k}, \quad g = \sum_{j=1}^{n} y_j \chi_{B_j}.$$

不妨设 $\sum_{k=1}^{m} A_k = \sum_{j=1}^{n} B_j = \Omega$. 于是

$$f + g = \sum_{k=1}^{m} \sum_{j=1}^{n} (x_k + y_j) \chi_{A_k \cap B_j},$$

从而

$$\begin{aligned}
\int (f + g) d\mu &= \sum_{k=1}^{m} \sum_{j=1}^{n} (x_k + y_j) \mu(A_k \cap B_j) \\
&= \sum_{k=1}^{m} \sum_{j=1}^{n} x_k \mu(A_k \cap B_j) + \sum_{k=1}^{m} \sum_{j=1}^{n} y_j \mu(A_k \cap B_j) \\
&= \sum_{k=1}^{m} x_k \mu(A_k) + \sum_{j=1}^{n} y_j \mu(B_j) \\
&= \int f d\mu + \int g d\mu.
\end{aligned}$$

其次设 f, g 是非负可测函数, 则存在非负简单函数序列 $\{f_n\}, \{g_n\}$ 使 $0 \leqslant f_n \uparrow f$,

$0 \leqslant g_n \uparrow g$, 因而 $0 \leqslant f_n + g_n \uparrow f + g$. 于是根据定义 2′

$$\begin{aligned}\int (f+g)d\mu &= \lim_{n\to\infty}\int (f_n+g_n)d\mu = \lim_{n\to\infty}\left(\int f_n d\mu + \int g_n d\mu\right)\\ &= \lim_{n\to\infty}\int f_n d\mu + \lim_{n\to\infty}\int g_n d\mu\\ &= \int f d\mu + \int g d\mu.\end{aligned}$$

再次设 f, g 是实可测函数, 由于 $\int f + \int g$ 存在, 则 $\int f^+ + \int g^+, \int f^- + \int g^-$ 至少有一有限, 不妨设后者有限. 对于任一非负可测函数 f, 若积分有限, 由定义 2 可知必然 a.e. 有限.$\left(\text{否则可选择 } s = \infty\chi_{\{f=\infty\}}, \text{ 则 } \int s d\mu = \infty \cdot \mu(f=\infty) = \infty \leqslant \int f d\mu.\right)$ 由此可知 f^-, g^- a.e. 有限. 因而 $f+g$ a.e. 有意义且

$$(f+g)^- \leqslant f^- + g^-, \text{a.e.},$$

再由

$$f+g = (f+g)^+ - (f+g)^- = (f^+ - f^-) + (g^+ - g^-), \text{及} (f+g)^-, f^-, g^-$$

的 a.e. 有限性可知

$$(f+g)^+ + f^- + g^- = f^+ + g^+ + (f+g)^-, \text{a.e.},$$

于是

$$\int (f+g)^+ + \int f^- + \int g^- = \int f^+ + \int g^+ + \int (f+g)^-,$$

再由 $\int f^-, \int g^-, \int (f+g)^-$ 的有限性可得

$$\int (f+g)^+ - \int (f+g)^- = \int f^+ - \int f^- + \int g^+ - \int g^-,$$

即

$$\int (f+g) = \int f + \int g.$$

至于 f, g 是复可测函数的情形, 分实部、虚部分别进行讨论, 可得同样结论.

b) 若 $\int f d\mu$ 存在, 则对任何 $A, B \in \mathscr{A}$, $\int_A f d\mu, \int_B f d\mu$ 存在且 $\int_A f + \int_B f$ 存在. 当 $A \cap B = \varnothing$ 时, 则

$$f\chi_{A+B} = f\chi_A + f\chi_B,$$

由 a) 知

$$\int_{A+B} f = \int f\chi_{A+B} = \int f\chi_A + \int f\chi_B = \int_A f + \int_B f.$$

c) 对 f 是非负简单函数, $c \geqslant 0$ 的情形即性质 2. 若 $c < 0$, 则由实函数积分的定义及 $(cf)^+ = 0, (cf)^- = -cf$ 知

$$\begin{aligned}\int cf d\mu &= \int (cf)^+ d\mu - \int (cf)^- d\mu \\ &= -\int (-c) f d\mu = -(-c)\int f d\mu = c\int f d\mu,\end{aligned}$$

其中第三个等式是由于 $-c > 0$ 及性质 2.

若 f 是非负可测函数, 当 $c \geqslant 0$ 时, 设 $0 \leqslant f_n \uparrow f, \{f_n\}$ 是简单函数序列, 则 $0 \leqslant cf_n \uparrow cf, \{cf_n\}$ 也是简单函数序列, 从而

$$\int cf d\mu = \lim_{n\to\infty}\int cf_n d\mu = c\lim_{n\to\infty}\int f_n d\mu = c\int f d\mu.$$

若 $c < 0$, 同 f 是非负简单函数的情形同样可证.

若 f 是实函数, c 有限, 则当 $c \geqslant 0$ 时

$$\begin{aligned}\int cf d\mu &= \int (cf)^+ d\mu - \int (cf)^- d\mu \\ &= \int cf^+ d\mu - \int cf^- d\mu \\ &= c\int f d\mu.\end{aligned}$$

而当 $c < 0$ 时, $(cf)^+ = -cf^-, (cf)^- = -cf^+$, 同样有

$$\int cf d\mu = c\int f d\mu.$$

对于 f 是实函数, c 是复数的情形, 若令 $c = c_1 + ic_2, c_1, c_2$ 是实数, 易知

$$\begin{aligned}\int cf d\mu &= \int c_1 f d\mu + i\int c_2 f d\mu \\ &= c_1\int f d\mu + ic_2\int f d\mu = c\int f d\mu.\end{aligned}$$

若 f 是复可测函数且 f 可积, 设

$$f = f_1 + if_2, \qquad c = c_1 + ic_2,$$

则 $cf = c_1f_1 - c_2f_2 + i(c_1f_2 + c_2f_1)$, 故

$$\int cf = \int (c_1f_1 - c_2f_2) + i\int (c_1f_2 + c_2f_1),$$

由于 f_1, f_2 可积, c_1, c_2 有限. 本定理 1)a) 的条件满足, 故上式

$$\begin{aligned}
&= \int c_1f_1 + \int (-c_2f_2) + i\left(\int c_1f_2 + \int c_2f_1\right) \\
&= c_1\int f_1 - c_2\int f_2 + ic_1\int f_2 + ic_2\int f_1 \\
&= c_1\int f_1 + ic_1\int f_2 + ic_2\int f_1 + (ic_2)i\int f_2 \\
&= c_1\int (f_1 + if_2) + ic_2\int (f_1 + if_2) \\
&= (c_1 + ic_2)\int (f_1 + if_2) = c\int f.
\end{aligned}$$

d) $\int \overline{f} = \overline{\int f}$, 由 c) 立刻得到.

2) a) 对于 f, g 是非负可测函数的情形已证 (性质 1′), 今证 f, g 是实函数的情形. 由 $f \geqslant g$ 知

$$f^+ \geqslant g^+, \qquad f^- \leqslant g^-,$$

从而对一切 $A \in \mathscr{A}$

$$\int_A f^+ + \int_A g^- \geqslant \int_A g^+ + \int_A f^-.$$

由于 $\int f, \int g$ 存在, 故若 $\int_A f^+$ 及 $\int_A g^-$ 中至少有一有限, 不妨设 $\int_A g^-$ 有限, 于是 $\int_A f^-$ 也有限, 就有

$$\int_A f^+ - \int_A f^- \geqslant \int_A g^+ - \int_A g^-;$$

若 $\int_A f^+ = +\infty, \int_A g^- = +\infty$, 则 $\int_A f = +\infty, \int_A g = -\infty$, 总之, 对一切 $A \in \mathscr{A}$, 有

$$\int_A f \geqslant \int_A g.$$

反之, 若 μ 是 σ-有限测度且对一切 $A \in \mathscr{A}$, $\int_A f d\mu \geqslant \int_A g d\mu$, 往证 $f \geqslant g$, a.e.. 即 $\mu(f < g) = 0$ 如若不然, 令

$$A_n = \left\{f < g - \frac{1}{n}\right\} \cap \{|g| < +\infty\},$$

$$B_m = \{f < m\} \cap \{g = +\infty\}.$$

则

$$\{f<g\}=\left(\bigcup_{n=1}^{\infty}A_n\right)\cup\left(\bigcup_{m=1}^{\infty}B_m\right).$$

若 $\mu\{f<g\}\neq 0$, 必有一 A_n 或 B_m 使 $\mu(A_n)\neq 0$ 或 $\mu(B_m)\neq 0$.不妨设 $\mu(A_n)\neq 0$. 再由 μ 的 σ-有限性, 存在集序列 $\{\Omega_k\}$, Ω_k 两两不交, $\mu(\Omega_k)$ 有限, $\sum\limits_{k=1}^{\infty}\Omega_k=\Omega$, 故必存在一 Ω_k 使 $0<\mu(A_n\cap\Omega_k)<\infty$, 于是取 $A=A_n\cap\Omega_k$, 有

$$\int_A fd\mu\leqslant\int_A\left(g-\frac{1}{n}\right)d\mu=\int_A gd\mu-\frac{1}{n}\mu(A)<\int_A gd\mu.$$

这与假设矛盾. 因而 2)a) 获证.

b) 当 f 是实函数时, 由于 $f\leqslant|f|,-f\leqslant|f|$, 故有

$$\pm\int f\leqslant\int|f|,$$

即

$$\left|\int f\right|\leqslant\int|f|.$$

当 f 是复函数时, 若 $\int|f|=\infty$, 不等式自然成立. 若 $\int|f|<\infty$, 由于 $\pm f_i\leqslant|f|, i=1,2$. 其中 f_1,f_2 分别为 f 的实部和虚部. 故 $\int f_i, i=1,2$ 有限. 即 $\int fd\mu$ 有限. 设

$$\int fd\mu=re^{it},r,t\text{实数},r=\left|\int fd\mu\right|.$$

$$f(\omega)=\rho(\omega)e^{i\alpha(\omega)},\rho(\omega)=|f(\omega)|,\alpha(\omega)\text{为}f(\omega)\text{的幅角}.$$

(当 $\rho(\omega)=0$ 时取 $\alpha(\omega)=0$), 易证 $\rho(\omega),\alpha(\omega)$ 均为可测函数.

$$\begin{aligned}\left|\int fd\mu\right|&=r=e^{-it}\int fd\mu\\&=e^{-it}\int\rho(\omega)e^{i\alpha(\omega)}d\mu\\&=\int\rho(\omega)e^{i(\alpha(\omega)-t)}d\mu\end{aligned}$$

$$= \int \rho(\omega)\cos(\alpha(\omega)-t)d\mu + i\int \rho(\omega)\sin(\alpha(\omega)-t)d\mu$$
$$= \int \rho(\omega)\cos(\alpha(\omega)-t)d\mu$$
$$\leqslant \int \rho(\omega)d\mu = \int |f|d\mu.$$

c) 若 $f \geqslant 0$, 由性质 4 知若 $f=0$, a.e., 则 $\int f = 0$. 反之, 若 $f \geqslant 0$, $\int f = 0$, 则必有 $f=0$, a.e.. 事实上, $\{f>0\} = \bigcup_{n=1}^{\infty}\left\{f > \frac{1}{n}\right\}$, 若 $\mu(f>0)>0$, 则由测度的连续性必存在 n, 使 $\mu\left(f>\frac{1}{n}\right)>0$, 令 $s = \frac{1}{n}\cdot \chi_{\{f>1/n\}}$, 则 $0 \leqslant s \leqslant f$, 因而

$$\int f d\mu \geqslant \int s d\mu = \frac{1}{n}\mu\left(f>\frac{1}{n}\right) > 0.$$

3) a) 若 $\int |f| < \infty$, 已由 2)a), 2)b) 可证 f 可积. 今设 $f = f_1 + if_2$ 可积, 则 $\int f_1, \int f_2$ 有限, 即 $\int f_1^{\pm}, \int f_2^{\pm}$ 有限. 而

$$|f| \leqslant |f_1| + |f_2| = f_1^+ + f_1^- + f_2^+ + f_2^-,$$

故由积分的线性性质知

$$\int |f| \leqslant \int f_1^+ + \int f_1^- + \int f_2^+ + \int f_2^- < \infty.$$

即 $|f|$ 可积.

至于由 $|f|$ 可积得知 f a.e. 有限, 在 1)a) 关于实可测函数的证明中已证.

b) $|f| \leqslant g$ 而 g 可积, 由 2)a) 知 $\int |f| \leqslant \int g < \infty$, 故 $|f|$ 可积, 再由 3)a) 知 f 可积.

c) f,g 可积, 则由 a) 知 f,g a.e. 有限. 故 $f+g$ a.e. 有意义且 $|f+g| \leqslant |f|+|g|$. 而

$$\int (|f|+|g|) = \int |f| + \int |g| < \infty.$$

故 $|f|+|g|$ 可积, 再由 b) 知 $f+g$ 可积.

d) 不难证明, 当 $\int |f|^2, \int |g|^2$ 至少有一个为 $+\infty$ 时, 不等式成立. 现在, 假设

$\int |f|^2, \int |g|^2$ 有限 (这时 $\int |fg|$ 有限). 对任意实数 a 和 b,

$$a^2 \int |f|^2 + 2ab \int |fg| + b^2 \int |g|^2 = \int (a|f| + b|g|)^2 \geqslant 0.$$

这说明上式左端是一非负定二次型, 由此知

$$\left(\int |fg| \right)^2 \leqslant \int |f|^2 \cdot \int |g|^2.$$

再利用 2)b),

$$\left| \int (fg) \right| \leqslant \int |fg|,$$

即得

$$\left| \int fg \right|^2 \leqslant \int |f|^2 \int |g|^2.$$

此不等式称为 Schwarz 不等式.

推论 若 f 是 A 上取非负值的可测函数，则对任何 $c > 0$ 有

$$\mu(\{f \geqslant c\} \cap A) \leqslant \frac{1}{c} \int_A f d\mu.$$

证 令 $A' = \{f \geqslant c\} \cap A$.

$$\int_A f d\mu = \int_{A'} f d\mu + \int_{A-A'} f d\mu \geqslant \int_{A'} f d\mu \geqslant c\mu(A').$$

此即所要证的不等式. 它是Чебышев不等式的推广.

习题及补充

1. 设 μ_1, μ_2 是可测空间 $(\Omega, \mathscr{A})$ 上的测度, a_1, a_2 是两个非负有限数, $\mu = a_1\mu_1 + a_2\mu_2$. 试证: 若 f 关于 μ_1, μ_2 的积分存在且 $a_1 \int f d\mu_1 + a_2 \int f d\mu_2$ 有意义, 则 f 关于 μ 的积分也存在, 且

$$\int f d\mu = a_1 \int f d\mu_1 + a_2 \int f d\mu_2.$$

2.(积分中值定理) 设 f, g 是测度空间 $(\Omega, \mathscr{A}, \mu)$ 上的可测函数, g 关于 μ 可积, $-\infty < a \leqslant f \leqslant b < +\infty$, a.e., 则存在一个常数 $c \in [a, b]$, 使

$$\int_\Omega f|g| d\mu = c \int_\Omega |g| d\mu.$$

特别, 若 μ 有限, 则 $\int fd\mu = c\mu(\Omega)$.

3. 设 f, g 是 $(\Omega, \mathscr{A}, \mu)$ 上取有限值的简单函数, 若 f, g 之一关于 μ 可积, 则 fg 也可积.

4. 设 f 是 $(\Omega, \mathscr{A}, \mu)$ 上的可积函数, 则对每一正数 $\varepsilon, \mu(\{\omega : |f(\omega)| \geqslant \varepsilon\}) < \infty$.

5. 设 Ω 是全体正整数组成的空间, $\mathscr{A}$ 是 Ω 的一切子集作成的 σ-代数, 对于 $A \in \mathscr{A}, \mu(A) = A$ 中点的个数, 则 $(\Omega, \mathscr{A}, \mu)$ 是一个测度空间. 讨论在此空间上函数的积分存在与可积的充分必要条件.

6. 若 f, g 是 $(\Omega, \mathscr{A})$ 上的可测函数. 用 $P(\mathscr{A})$ 表示 $\mathscr{A}$ 上的一切概率测度的集, 若对一切 $\nu \in P(\mathscr{A}), \int fd\nu = \int gd\nu$, 则 $f(\omega) = g(\omega)$, 对一切 $\omega \in \Omega$ 成立.

§3.3　收敛定理

这一节我们将讨论积分号下的极限运算, 几个著名的收敛定理在测度论及概率论中起着重要的作用.

§2 证明的单调收敛定理是其中重要的一个. 现在我们再来证明另外两个收敛定理, 即 Fatou-Lebesgue 定理及控制收敛定理以及它们的推论.

定理 1(Fatou-Lebesgue 定理)　设 g, h 是实值可积函数, $\{f_n\}$ 是实可测函数序列,

i) 若对一切 $n \geqslant 1, g \leqslant f_n$, 则

$$\int \varliminf_{n\to\infty} f_n \leqslant \varliminf_{n\to\infty} \int f_n; \tag{1}$$

ii) 若对一切 $n \geqslant 1, f_n \leqslant h$, 则

$$\varlimsup_{n\to\infty} \int f_n \leqslant \int \varlimsup_{n\to\infty} f_n; \tag{2}$$

iii) 若 $g \leqslant f_n \uparrow f$ 或对一切 $n \geqslant 1, g \leqslant f_n \leqslant h, f_n \xrightarrow{\text{a.e.}} f$, 则 $\int f_n \to \int f$.

证　首先, 由题设条件易知对一切 $n \geqslant 1, \int f_n d\mu$ 存在. 先证

i) 在 $g = 0$ 时成立, 这时 $0 \leqslant f_n$, 令 $g_n = \inf\limits_{k\geqslant n} f_k$, 易见

$$0 \leqslant g_n \uparrow \sup_n \inf_{k\geqslant n} f_k = \varliminf_{n\to\infty} f_n.$$

根据单调收敛定理

$$\int \varliminf_{n\to\infty} f_n = \lim_{n\to\infty} \int g_n.$$

同时 $g_n \leqslant f_n$, 所以 $\int g_n \leqslant \int f_n$, 取下极限又得

$$\varliminf_{n\to\infty} \int g_n \leqslant \varliminf_{n\to\infty} f_n.$$

这就证明了在 $g=0$ 时 (1) 成立.

至于一般情况, 我们只需注意到 $0 \leqslant f_n - g$, a.e., 即可利用上面已证结果得出

$$\int \varliminf_{n\to\infty} f_n - \int g = \int \varliminf_{n\to\infty} (f_n - g) \leqslant \varliminf_{n\to\infty} \int (f_n - g)$$
$$= \varliminf_{n\to\infty} \int f_n - \int g.$$

由于 g 可积, $\int g$ 有限, 故可由上式得出

$$\int \varliminf_{n\to\infty} f_n \leqslant \varliminf_{n\to\infty} f_n.$$

ii) 由于 $\varlimsup\limits_{n\to\infty} f_n = -\varliminf\limits_{n\to\infty} (-f_n)$ 以及 $-h \leqslant -f_n, -h$ 可积, 不难由 i) 导出 (2) 成立.

iii) 当 $g \leqslant f_n \uparrow f$ 时, 显然 $0 \leqslant f_n - g \uparrow f - g$, a.e., 根据单调收敛定理,

$$\lim_{n\to\infty} \int f_n - \int g = \lim_{n\to\infty} \int (f_n - g) = \int (f - g) = \int f - \int g.$$

由于 $\int g$ 有限, 故得 $\lim\limits_{n\to\infty} \int f_n = \int f$.

最后, 当 $g \leqslant f_n \leqslant h, f_n \xrightarrow{\text{a.e.}} f$ 时, 设 $A = \{f_n \not\to f\}$, 则 $\mu(A) = 0$. 令

$$f_n^* \triangleq f_n \chi_{A^c} + g\chi_A; \quad f^* = f\chi_{A^c} + g\chi_A.$$

显然 $f_n^* = f_n$, a.e., $f^* = f$, a.e., 且 $g \leqslant f_n^* \leqslant h$, $\varlimsup\limits_{n\to\infty} f_n^* = \varliminf\limits_{n\to\infty} f_n^* = f^*$. 于是 $\{f_n^*\}$ 满足 i) 及 ii) 的条件, 因而

$$\int f^* \leqslant \varliminf_{n\to\infty} \int f_n^* \leqslant \varlimsup_{n\to\infty} \int f_n^* \leqslant \int f^*,$$

即

$$\int f^* = \lim_{n\to\infty} \int f_n^*.$$

但 $\int f_n^* = \int f_n, \int f^* = \int f$, 所以

$$\int f = \lim_{n\to\infty} \int f_n. \qquad \square$$

容易看到, 当定理中的条件几乎成立时, 定理的结论仍然成立.

定理 2(控制收敛定理)　设 g 是可积函数, $|f_n| \leqslant g$, a.e., 如果 $f_n \xrightarrow{\text{a.e.}} f$ 或 $f_n \xrightarrow{\mu} f$, 则 $\int |f_n - f| \to 0$, 因而 $\int f_n \to \int f$.

证　由于 $|f_n| \leqslant g$, a.e., $f_n \xrightarrow{\text{a.e.}} f$ 或 $f_n \xrightarrow{\mu} f$, 故知 $|f| \leqslant g$, a.e., 因此 f 可积, $\left|\int f_n - \int f\right| \leqslant \int |f_n - f|$, 只需证 $\int |f_n - f| \to 0$ 即知 $\int f_n \to \int f$.

当 $f_n \xrightarrow{\text{a.e.}} f$ 时, $|f_n - f| \xrightarrow{\text{a.e.}} 0$ 且 $|f_n - f| \leqslant 2g$, a.e., 由定理 1 iii) 知 $\int |f_n - f| \to 0$.

以下往证当 $f_n \xrightarrow{\mu} f$ 时 $\int |f_n - f| \to 0$.

事实上, 令 $g_n = |f_n - f|$, 由给定条件知 $0 \leqslant g_n \leqslant 2g$, a.e. 且 $g_n \xrightarrow{\mu} 0$. 首先存在一个子序列 $\{g_{n_k}\}$ 使

$$\int g_{n_k} \to \varlimsup_{n\to\infty} \int g_n,$$

这里 $g_{n_k} \xrightarrow{\mu} 0$, 所以又有一子序列 $g_{n_{k_i}} \xrightarrow{\text{a.e.}} 0$, 应用本定理已证结果知

$$\int g_{n_{k_i}} \to 0.$$

但 $\left\{\int g_{n_{k_i}}\right\}$ 是 $\left\{\int g_{n_k}\right\}$ 的子序列, 它们有相同的极限, 于是有

$$\varlimsup_{n\to\infty} \int g_n = 0.$$

但 $g_n \geqslant 0$, 因而 $\varliminf_{n\to\infty} \int g_n \geqslant 0$, 故由上述等式知 $\lim_{n\to\infty} \int g_n = 0$, 即

$$\lim_{n\to\infty} \int |f_n - f| = 0.$$

定理获证. □

推论 1　如果 $\{f_n\}$ 是可测函数序列, 若 $f_n, n \geqslant 1$, 非负或 $\sum_{n=1}^{\infty} \int |f_n| d\mu < \infty$, 则 $\sum_{n=1}^{\infty} f_n$ a.e. 收敛, 积分存在且

$$\int \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int f_n. \tag{3}$$

证　令 $g_n = \sum_{k=1}^{n} |f_k|, g = \sum_{k=1}^{\infty} |f_k|$, 则 $0 \leqslant g_n \uparrow g$, 利用单调收敛定理, 当 f_n 非

负时即得 (3) 式. 当 $\sum_{n=1}^{\infty}\int|f_n|<\infty$ 时, 知 g 可积, 因而 a.e. 有限, $\sum_{n=1}^{\infty}f_n$ a.e. 收敛且对一切 $n\geqslant 1$, $\left|\sum_{k=1}^{n}f_k\right|\leqslant g$, 应用控制收敛定理即得 (3) 式.

推论 2 若 $\int f d\mu$ 存在, 则对于任意的 $A=\sum_{n=1}^{\infty}A_n, A\in\mathscr{A}, A_n\in\mathscr{A}, n=1,2,\cdots$, 两两不交, 有

$$\int_A f d\mu=\sum_{k=1}^{\infty}\int_{A_k}f d\mu. \tag{4}$$

证 对于实可测函数 f, 我们有

$$f^{\pm}\chi_A=\sum_{k=1}^{\infty}f^{\pm}\chi_{A_k},$$

利用推论 1,

$$\int_A f^+ d\mu=\sum_{k=1}^{\infty}\int_{A_k}f^+ d\mu,$$

$$\int_A f^- d\mu=\sum_{k=1}^{\infty}\int_{A_k}f^- d\mu.$$

由于 f 的积分存在, 上面的两个级数中至少有一个收敛于有限和, 因此可以逐项相减, 从而 (4) 成立.

对于复值函数 f 来说, 可以通过考虑它的实部和虚部而得出 (4). □

当 f 的积分存在时, 它的积分

$$\varphi(A)=\int_A f d\mu, A\in\mathscr{A},$$

是 $\mathscr{A}$ 上的一个集函数, 称为 f 在 $\mathscr{A}$ 上关于 μ 的**不定积分**. 推论 2 说明, 在 Ω 上积分存在的可测函数 f 的不定积分是一 σ- 可加集函数.

推论 3 如果 f 可积, 则当 $\mu(A)\to 0$ 时

$$\int_A|f|d\mu\to 0. \tag{5}$$

证 令

$$f_n=|f|\chi_{\{|f|<n\}}+n\chi_{\{|f|\geqslant n\}},$$

显然 $0\geqslant f_n\uparrow|f|$, 由单调收敛定理

$$\lim_{n\to\infty}\int f_n=\int|f|.$$

由于 f 可积, 因而 $\int |f| < \infty$, 故对任一 $\varepsilon > 0$, 存在一自然数 n_0, 使得

$$\int (|f| - f_{n_0}) = \int |f| - \int f_{n_0} < \frac{\varepsilon}{2},$$

由此可见当 $\mu(A) < \dfrac{\varepsilon}{2n_0}$ 时

$$\int_A |f| = \int_A f_{n_0} + \int_A (|f| - f_{n_0}) \leqslant n_0 \mu(A) + \int (|f| - f_{n_0}) < \varepsilon.$$

即当 $\mu(A) \to 0$ 时, $\int_A |f| \to 0$. □

在单调收敛定理、Fatou-Lebesgue 定理及控制收敛定理中, 参数 $n \to \infty$ 都可以换成在任意集 $T \subset R$ 内取值的参数 $t \to t_0 (t_0 \in \tilde{R})$, 这是因为 $\lim\limits_{t \to t_0} a_t = a$ 的充分必要条件是对任意 $\{t_n\} \subset T, t_n \to t_0 (n \to \infty)$ $\lim\limits_{n \to \infty} a_{t_n} = a$. 例如, 控制收敛定理对连续参数情形的推广为

推论 4 若对一切 $t \in T_0$(当 t_0 有限时, 存在 $\delta > 0, T_0 = (t_0 - \delta, t_0 + \delta)$, 当 $t_0 = +\infty$ 时 $T_0 = (M, \infty)$, 当 $t_0 = -\infty$ 时 $T_0 = (-\infty, -M)$, 其中 M 是某一充分大的正数), $|f_t| \leqslant g$, a.e., g 关于 μ 可积, $\lim\limits_{t \to t_0} f_t = f_{t_0}$, a.e. 或 $f_t \xrightarrow{\mu} f_{t_0}$(当 $t \to t_0$), 则 $\int |f_t - f_{t_0}| d\mu \to 0 (t \to t_0)$, 因而有

$$\lim_{t \to t_0} \int f_t d\mu = \int f_{t_0} d\mu.$$

证明留作习题.

利用这一结果, 我们可以得到下列关于积分号下微分、积分的一系列推论.

推论 5 若对每一 $t \in T, f_t$ 关于 μ 可积, 而对每一 $\omega \in \Omega$,看作 t 的函数时沿集合 T 在 t_0 点有导数 $\left(\dfrac{df_t(\omega)}{dt}\right)_{t_0}$, 且 $\left|\dfrac{f_t(\omega) - f_{t_0(\omega)}}{t - t_0}\right| \leqslant g(\omega)$, g 可积, 则

$$\left(\frac{d}{dt} \int f_t(\omega) d\mu(\omega)\right)_{t_0} = \int \left(\frac{df_t(\omega)}{dt}\right)_{t_0} d\mu(\omega). \tag{6}$$

证 根据导数的定义

$$\frac{f_t(\omega) - f_{t_0}(\omega)}{t - t_0} \to \left(\frac{df_t(\omega)}{dt}\right)_{t_0}, t \to t_0, t, t_0 \in T.$$

而由假设知 g 可积, $\left|\dfrac{f_t(\omega)-f_{t_0}(\omega)}{t-t_0}\right|\leqslant g(\omega), t\in T, \omega\in\Omega$, 于是利用推论 4 知

$$\lim_{t\to t_0}\int\frac{f_t(\omega)-f_{t_0}(\omega)}{t-t_0}d\mu(\omega)=\int\left(\frac{df_t(\omega)}{dt}\right)_{t_0}d\mu(\omega).$$

而

$$\begin{aligned}&\lim_{t\to t_0}\int\frac{f_t(\omega)-f_{t_0}(\omega)}{t-t_0}d\mu(\omega)\\=&\lim_{t\to t_0}\frac{\displaystyle\int f_t(\omega)d\mu(\omega)-\int f_{t_0}(\omega)d\mu(\omega)}{t-t_0}\\=&\left(\frac{d}{dt}\int f_t(\omega)d\mu(\omega)\right)_{t_0},\end{aligned}$$

故得 (6) 式. □

推论 6 如果在区间 $[a,b]$ 上, $\int f_t(\omega)d\mu(\omega)$ 有限, 对一切 $\omega\in\Omega, t\in(a,b)$, $\dfrac{df_t(\omega)}{dt}$ 存在, 且 $\left|\dfrac{df_t(\omega)}{dt}\right|\leqslant g(\omega), g$ 可积, 则在 (a,b) 上

$$\frac{d}{dt}\int f_t(\omega)d\mu(\omega)=\int\frac{df_t(\omega)}{dt}d\mu(\omega). \tag{7}$$

证 当 f_t 是实函数时, 对每一 $\omega\in\Omega$, 利用 Lagrange 公式

$$f_t(\omega)-f_{t_0}(\omega)=(t-t_0)\left(\frac{df_t(\omega)}{dt}\right)_{t'}, t'\in(a,b),$$

所以对任意 $t_0\in(a,b), \omega\in\Omega$,

$$\left|\frac{f_t(\omega)-f_{t_0(\omega)}}{t-t_0}\right|=\left|\left(\frac{df_t(\omega)}{dt}\right)_{t'}\right|\leqslant g(\omega).$$

由推论 5 即得 (7) 式. 当 f_t 是复函数时, 分别对实部和虚部进行同样的证明, 可知 (7) 式成立. □

推论 7 如果在一有限区间 $[a,b]$ 上, 对一切 $\omega\in\Omega, f_t(\omega)$ 对 t 连续, 且对一切 $t\in[a,b], \omega\in\Omega, |f_t(\omega)|\leqslant g(\omega), g$ 可积, 则对一切 $t\in[a,b]$,

$$\int_a^t\left(\int f_\tau(\omega)d\mu(\omega)\right)d\tau=\int\left(\int_a^t f_\tau(\omega)d\tau\right)d\mu(\omega). \tag{8}$$

又若以上的假定在每一有限区间上都成立, 且 $\int_{-\infty}^{\infty}|f_t(\omega)|dt \leqslant h(\omega)$, h 关于 μ 可积, 则

$$\int_{-\infty}^{\infty}\int f_t(\omega)d\mu(\omega)dt = \int\left(\int_{-\infty}^{\infty} f_t(\omega)dt\right)d\mu(\omega). \tag{9}$$

其中关于 t 的积分都是黎曼积分.

证　(8) 式的左端是 $\int f_t(\omega)d\mu(\omega)$ 的原函数 $\Big($由控制收敛定理易知 $\int f_t(\omega)d\mu(\omega)$ 是关于 t 的连续函数, 因而原函数存在$\Big)$, 由于 (8) 式两端在 $t=a$ 时皆为零, 因此只需证 (8) 式右端也是 $\int f_t(\omega)d\mu(\omega)$ 的原函数, 即右端关于 t 的导数等于 $\int f_t(\omega)d\mu(\omega)$.

由于 $\left|\dfrac{d}{dt}\int_a^t f_\tau(\omega)d\tau\right| = |f_t(\omega)| \leqslant g(\omega)$, g 可积, 同时 $\left|\int_a^t f_\tau(\omega)d\tau\right| \leqslant \int_a^t |f_\tau(\omega)|d\tau \leqslant (b-a)g(\omega)$, 即 $\int_a^t f_\tau(\omega)d\tau$ 关于 μ 可积, 利用推论 6 即得

$$\begin{aligned}\frac{d}{dt}\int\left(\int_a^t f_\tau(\omega)d\tau\right)d\mu(\omega) &= \int\frac{d}{dt}\left(\int_a^t f_\tau(\omega)d\tau\right)d\mu(\omega)\\ &= \int f_t(\omega)d\mu(\omega).\end{aligned}$$

所以 (8) 式右端是 $\int f_t(\omega)d\mu(\omega)$ 的原函数, 于是对任何 $t\in[a,b]$,(8)式成立.

$\left(\int_a^t f_\tau(\omega)d\tau\right)$ 的可测性可用第 4 章的 Fubini 定理证明. 根据前一部分的假设, 对于任意 $a\leqslant b$, 有

$$\int_a^b\left(\int f_t(\omega)d\mu(\omega)\right)dt = \int\left(\int_a^b f_t(\omega)dt\right)d\mu(\omega),$$

$$\left|\int_a^b f_t(\omega)dt\right| \leqslant \int_a^b |f_t(\omega)|dt \leqslant \int_{-\infty}^{\infty}|f_t(\omega)|dt \leqslant h(\omega),$$

h 关于 μ 可积. 利用控制收敛定理, 令 $a\to-\infty, b\to+\infty$ 即得

$$\int_{-\infty}^{\infty}\left(\int f_t(\omega)d\mu(\omega)\right)dt = \int\left(\int_{-\infty}^{\infty} f_t(\omega)dt\right)d\mu(\omega).$$ □

对于无穷项级数的问题, 也可以利用收敛定理得到一系列关于求和次序互换、逐项取极限、逐项微分、逐项积分的条件. 事实上, 我们令 $\Omega=\{1,2,\cdots\}$, $\mathscr{A}=\Omega$

的一切子集类, 令 $\mu(\{n\})=1, n=1,2,\cdots$ 则 $(\Omega,\mathscr{A},\mu)$ 是一个 σ-有限测度空间 (称此测度 μ 为计数测度), $\int f d\mu=\sum_{n=1}^{\infty} f_n$, f 关于 μ 可积的充分必要条件是级数绝对收敛.

推论 8 $f_{nm}\geqslant 0$, 或 $\sum_{k=1}^{m}|f_{nk}|\leqslant g_n$, 对一切 $m,n\geqslant 1$ 成立, $\sum_{n=1}^{\infty} g_n<\infty$, 则

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} f_{nm}=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} f_{nm}. \tag{10}$$

推论 9 当 $m\to\infty$ 时, $0\leqslant f_{nm}\uparrow f_n$ 或 $|f_{nm}|\leqslant g_n$, $\sum_{m=1}^{\infty} g_n<\infty$ 且 $\lim_{n\to\infty} f_{nm}=f_n$ 对一切 $n\geqslant 1$ 成立, 则

$$\lim_{m\to\infty}\sum_{n=1}^{\infty} f_{nm}=\sum_{n=1}^{\infty} f_n. \tag{11}$$

推论 10 若 $0\leqslant f_{nm}$, 则

$$\varliminf_{m\to\infty}\sum_{n=1}^{\infty} f_{nm}\geqslant\sum_{n=1}^{\infty}\varliminf_{m\to\infty} f_{nm}.$$

这些推论的证明以及关于逐项微分、逐项积分的命题的叙述和证明留给读者.

习题及补充

1. 利用 Fatou-Lebesgue 定理证明: 若 μ 是测度,

i) 对任意可测集列 $\{A_n\}$ 有

$$\mu(\varliminf_{n\to\infty} A_n)\leqslant\varliminf_{n\to\infty}\mu(A_n).$$

ii) 若 μ 有穷, 则

$$\mu(\varlimsup_{n\to\infty} A_n)\geqslant\varlimsup_{n\to\infty}\mu(A_n).$$

2. 若对一切 $t\in(t_0-\delta,t_0+\delta)$, $|f_t(\omega)|\leqslant g(\omega)$ 除零测集以外成立, g 关于 μ 可积且 $\lim_{t\to t_0} f_t=f_{t_0}$, a.e. 或 $f_t\xrightarrow{\mu} f_{t_0}$(当 $t\to t_0$), 则

$$\lim_{t\to t_0}\int f_t(\omega)d\mu(\omega)=\int f_{t_0}(\omega)d\mu(\omega).$$

3. 利用控制收敛定理证明

$$\lim_{t\to 0+}\sum_{n=1}^{\infty}\frac{\cos nt}{2^n+t^n}=\sum_{n=1}^{\infty}\frac{1}{2^n}.$$

4. f 是 $(\Omega,\mathscr{A},\mu)$ 上的可测函数, 若 f 可积, 则集 $\{\omega\colon f(\omega)\neq 0\}$ 可表示为可数个 $\mathscr{A}$ 可测且测度有限的集合的不交并, 如果仅是 $\int f d\mu$ 存在呢? $\left(\text{提示: }\int |f|d\mu = \sum_{n=1}^{\infty}\int_{\{n\leqslant|f|<n+1\}}|f|d\mu + \sum\int_{\frac{1}{n+1}\leqslant|f|<\frac{1}{n}}|f|d\mu.\right)$

5. 设 $\Omega=\{1,2,\cdots\}$, $\mathscr{A}=\Omega$ 的一切子集类, μ 为计数测度, $f(\omega)=0,\omega\in\Omega$,

$$f_n(\omega)=\begin{cases}\dfrac{1}{n}, & \omega\in\{1,\cdots,n\},\\ 0, & \omega\notin\{1,\cdots,n\}\end{cases}\quad n=1,2,\cdots.$$

则 $f_n(\omega)\to f(\omega)$ 对一切 ω 一致成立. 能否由此得出 $\lim\limits_{n\to\infty}\int f_n(\omega)d\mu=\int f(\omega)d\mu$?

6. 若 $\int f d\mu$ 有限, 则对一切 $A\in\mathscr{A}$, $\int_A f_n d\mu$ 一致地趋于 $\int_A f d\mu$ 的充分必要条件是 $\int |f_n-f|d\mu\to 0$.

7. 如果 $0\leqslant f_n\xrightarrow{\mu}f$(或 $0\leqslant f_n\xrightarrow{\text{a.e.}}f$), 则当有限的 $\int f_n\to\int f$(有限) 时, $\int_A f_n\to\int_A f$ 对于 $A\in\mathscr{A}$ 一致成立.(提示:$0\leqslant(f-f_n)^+\leqslant$ 可积函数 f, 且

$$\int(f-f_n)^+-\int(f-f_n)^-\to 0.)$$

§3.4　随机变量函数的数学期望的 L–S 积分表示与积分变换定理

在这一节, 我们首先给出随机变量的数学期望、方差、矩等数字特征的定义. 我们会看到随机变量的方差、矩以及更广一些随机变量函数的分布律都可以归结为随机变量函数的数学期望. 为了统一解决这些问题, 利用积分变换定理给出随机变量函数的数学期望的 L–S 积分表示, 并把它们转换到 $(R^{(n)},\mathscr{B}^{(n)},P_\xi)$ 空间上进行计算.

3.4.1　数学期望、方差及矩

定义 1　设 ξ 是概率场 $(\Omega,\mathscr{A},P)$ 上的一个随机变量 (实或复值), 如果 ξ 在 Ω 上关于 P 的积分存在, 则 ξ 在 Ω 上的积分称为 ξ 的数学期望. 记作

$$E\xi=\int\xi(\omega)dP.\tag{1}$$

这个概念比一般初等概率的概念稍广一些, 一般初等概率论书中要求 ξ 可积, 即 $E|\xi|$ 有限.

根据一般积分的性质, 立刻可以得到数学期望的性质.

性质 1(可加性) 如果概率场 $(\Omega,\mathscr{A},P)$ 上的随机变量 $\xi_1,\cdots,\xi_n$ 的数学期望都存在, 且 $E\xi_1+\cdots+E\xi_n$ 有意义, 则 $\xi_1+\cdots+\xi_n$ 的数学期望存在且

$$E(\xi_1+\cdots+\xi_n)=E\xi_1+\cdots+E\xi_n. \tag{2}$$

性质 2 设 c 是一有限常数, ξ 的数学期望存在, 则 $c\xi$ 的数学期望存在且

$$E(c\xi)=cE\xi. \tag{3}$$

如果 ξ 是复随机变量, 则要求 $E|\xi|$ 有限.

性质 3(乘法定理) 如果概率场 $(\Omega,\mathscr{A},P)$ 上的随机变量 $\xi_1,\cdots,\xi_n$ 独立且非负或有有限的数学期望, 则

$$E(\xi_1\cdots\xi_n)=E\xi_1\cdots E\xi_n. \tag{4}$$

证 只需证明 $n=2$ 的情形, 对一般情形不难用数学归纳法推出.

i) 设 ξ,η 是非负可测函数的情形已获证明, 当 ξ,η 是一般实可积函数 (数学期望有限) 时, 由

$\xi\eta=(\xi^+-\xi^-)(\eta^+-\eta^-)=\xi^+\eta^++\xi^-\eta^--\xi^-\eta^+-\xi^+\eta^-$ 等式右端积分存在且有限 (可积) 因而 $\xi\eta$ 也可积, 注意到 $\{\xi^+,\xi^-\}$ 与 $\{\eta^+,\eta^-\}$ 的独立性,

$$\begin{aligned}E\xi\eta&=E\xi^+\eta^++E\xi^-\eta^--E\xi^-\eta^+-E\xi^+\eta^-\\&=E\xi^+E\eta^++E\xi^-\eta^--E\xi^-\eta^+-E\xi^+\eta^-\\&=E(\xi^+-\xi^-)E(\eta^+-\eta^-)=E\xi E\eta.\end{aligned}$$

于是只需对 ξ,η 是非负可测函数的情形证明本性质. 下面我们用函数形式单调类定理证明该性质.

ii) 设 ξ,η 独立, 非负可测, η 是非负简单函数, 不妨设 $\eta=\sum\limits_{k=1}^{m}a_k\chi_{B_k}$, $\mathscr{L}$ 为 Ω 上实函数集.

$L\hat{=}\{\xi\geqslant 0,$ 可测, 与η 独立, $E\xi\eta=E\xi E\eta\}$ 则 L 为 $\mathscr{L}$-系. 事实上,

I) $1\in L$, 显然;

II) L 中有限个函数的线性组合仍属于 L;

III) 如果 $\xi_n\in L, 0\leqslant\xi_n\uparrow\xi$, 则由单调收敛定理及 η 非负可知 $E\xi\eta=\lim\limits_{n\to\infty}E\xi_n\eta=\lim E\xi_nE\eta=E\xi E\eta$, 且作为与 η 独立的函数的极限, ξ 与 η 独立. 当 $\xi=\chi_A$ 为示性

函数时, 由 ξ,η 独立, 知集合 A 与 $\{B_1,B_2,\cdots,B_m\}$ 独立, 因此

$$\begin{aligned}E\xi\eta &= E\sum_{k=1}^{m}a_k\chi_A\chi_{B_k}=\sum_{k=1}^{m}a_kP(AB_k)=\sum_{k=1}^{m}a_kP(A)P(B_k)\\ &=P(A)\sum_{k=1}^{m}a_kP(B_k)=E\xi E\eta,\end{aligned}$$

即 L 包含一切可测集的示性函数. 由单调类定理可知 L 包含一切与 η 独立的非负随机变量.

iii) 当 ξ,η 均为非负可测, ξ,η 独立时, 令

$$\eta_n(\omega)=\begin{cases}\dfrac{k}{2^n}, & \omega\in B_k=\left\{\dfrac{k}{2^n}\leqslant\eta<\dfrac{k+1}{2^n}\right\}, k=0,1,\cdots,n2^n-1,\\ n, & \omega\in B_{n2^n}=\{\eta\geqslant n\}.\end{cases}$$

此时 ξ,η_n 仍独立, 由 ii) 及单调收敛定理知

$$E\xi\eta=\lim_{n\to\infty}E\xi\eta_n=\lim_{n\to\infty}E\xi E\eta_n=E\xi E\eta.$$

显然, 如果 $E\xi,E\eta$ 有限, 则 $E\xi\eta$ 有限.

iv) 最后设 ξ,η 是两个独立的复随机变量, 已知它们分别有有限的数学期望, 仿 i) 可以推出 $|\xi\eta|$ 可积, 设 $\xi=\xi_1+i\xi_2,\eta=\eta_1+i\eta_2$, 其中 $\xi_1,\xi_2,\eta_1,\eta_2$ 是实随机变量. 根据复随机变量独立性定义, $(\xi_1,\xi_2),(\eta_1,\eta_2)$ 独立, 且它们的数学期望都有限, 于是利用 i) 得到

$$\begin{aligned}E\xi\eta &= E(\xi_1+i\xi_2)(\eta_1+i\eta_2)\\ &=E\xi_1\eta_1+iE\xi_1\eta_2+iE\xi_2\eta_1-E\xi_2\eta_2\\ &=E\xi_1E\eta_1+iE\xi_1E\eta_2+iE\xi_2E\eta_1-E\xi_2E\eta_2\\ &=(E\xi_1+iE\xi_2)(E\eta_1+iE\eta_2)\\ &=E\xi\cdot E\eta.\end{aligned}$$

□

例 1　设 $(\xi_1,\cdots,\xi_n)$ 是概率场 $(\Omega,\mathscr{A},P)$ 上的 n 维实随机变量, $t_1,\cdots,t_n$ 是实参数, 则随机变量 $\eta=e^{i(t_1\xi_1+\cdots+t_n\xi_n)}$ 的数学期望有限. 事实上

$$E|\eta|=1\cdot P(\Omega)=1.$$

故对任何 $(t_1,\cdots,t_n)\in R^{(n)}$,

$$f(t_1,\cdots,t_n)=Ee^{i(t_1\xi_1+\cdots+t_n\xi_n)}\tag{5}$$

有意义, 称为 $(\xi_1,\cdots,\xi_n)$ 的**特征函数**. 它和分布函数一样, 可以完全刻划随机变量的概率特性, 是一个很重要的工具, 我们将在第 6 章专门讨论.

特别当 $\xi_1,\cdots,\xi_n$ 独立时, $e^{it_1\xi_1},\cdots,e^{it_n\xi_n}$ 也独立, 于是得

$$\begin{aligned} f(t_1,\cdots,t_n) &= Ee^{it_1\xi_1}\cdots Ee^{it_n\xi_n} \\ &= f_1(t_1)\cdots f_n(t_n), \end{aligned} \tag{6}$$

其中 $f_k(t_k)$ 是 ξ_k 的特征函数, $k=1,\cdots,n$.

定义 2 设 ξ 是概率场 $(\Omega,\mathscr{A},P)$ 上的随机变量, ξ 几乎处处有限, $E\xi$ 存在, 则 $E|\xi-E\xi|^2$ 称为 ξ 的**方差**, 记作

$$D\xi = E|\xi-E\xi|^2. \tag{7}$$

当 $E\xi$ 有限时, 由于

$$\begin{aligned} |\xi-E\xi|^2 &= (\xi-E\xi)(\overline{\xi}-\overline{E\xi}) \\ &= |\xi|^2-\xi\cdot\overline{E\xi}-\overline{\xi}\cdot E\xi+|E\xi|^2, \end{aligned}$$

利用数学期望的性质, 立得

$$D\xi = E|\xi|^2-|E\xi|^2. \tag{8}$$

性质 4 如果概率场 $(\Omega,\mathscr{A},P)$ 上的随机变量 $\xi_1,\cdots,\xi_n$ 分别有有限的数学期望, 且两两独立, 则

$$D(\xi_1+\cdots+\xi_n) = D\xi_1+\cdots+D\xi_n. \tag{9}$$

特别当 $\xi_1,\cdots,\xi_n$ 有有限方差时, $\xi_1+\cdots+\xi_n$ 也有有限方差.

证 因为 $\xi_1,\cdots,\xi_n$ 有有限数学期望, 利用性质 1, $\xi_1+\cdots+\xi_n$ 有有限数学期望且

$$E(\xi_1+\cdots+\xi_n) = E\xi_1+\cdots+E\xi_n.$$

由此有

$$\begin{aligned} &|(\xi_1+\cdots+\xi_n)-E(\xi_1+\cdots+\xi_n)|^2 \\ &= [(\xi_1-E\xi_1)+\cdots+(\xi_n-E\xi_n)][\overline{(\xi_1-E\xi_1)}+\cdots+\overline{(\xi_n-E\xi_n)}] \\ &= |\xi_1-E\xi_1|^2+\cdots+|\xi_n-E\xi_n|^2+\sum_{i\neq j}(\xi_i-E\xi_i)\overline{(\xi_j-E\xi_j)}. \end{aligned}$$

由于 ξ_i 与 $\xi_j(i \neq j)$ 独立, 因而 $\xi_i - E\xi_i$ 与 $\overline{\xi_j - E\xi_j}$ 独立且有有限的数学期望, 于是当 $i \neq j$ 时,

$$E(\xi_i - E\xi_i)\overline{(\xi_j - E\xi_j)} = E(\xi_i - E\xi_i)E\overline{(\xi_j - E\xi_j)} = 0.$$

从而得到

$$\begin{aligned} &E|(\xi_1 + \cdots + \xi_n) - E(\xi_1 + \cdots + \xi_n)|^2 \\ &= E|\xi_1 - E\xi_1|^2 + \cdots + E|\xi_n - E\xi_n|^2. \end{aligned}$$

此即 (9) 式. □

定义 3　设 ξ, η 是概率场 $(\Omega, \mathscr{A}, P)$ 上的随机变量, 数学期望都存在, 则

$$b_{\xi\eta} \hat{=} E(\xi - E\xi)\overline{(\eta - E\eta)}, \tag{10}$$

(如果存在) 称为 ξ 对 η 的**相关矩**, 又若 $D\xi, D\eta \neq 0$ 且有限, 则

$$r_{\xi\eta} = \frac{b_{\xi\eta}}{\sqrt{D\xi D\eta}} \tag{11}$$

称为 ξ, η 的**相关系数**.

当 $\xi = \eta$ 时, $b_{\xi\eta} = D\xi$.

性质 5　若 $b_{\xi\eta}, r_{\xi\eta}$ 存在, 则

i) $b_{\xi\eta} = \overline{b_{\eta\xi}}, r_{\xi\eta} = \overline{r_{\eta\xi}}$;

ii) $|b_{\xi\eta}| \leqslant \sqrt{D\xi \cdot D\eta}, |r_{\xi\eta}| \leqslant 1$, 且若 $D\xi \neq 0$, 则等号成立的充分必要条件是

$$P(\eta = a\xi + b) = 1,$$

其中

$$a = \overline{b_{\xi\eta}}/D\xi, \qquad b = E\eta - aE\xi.$$

iii) 若 ξ, η 独立且都有有限的数学期望, 则 $b_{\xi\eta} = 0, r_{\xi\eta} = 0$.

证　i), iii) 显然, ii) 的第一部分由积分的性质立刻推出, 今证 ii) 的后一部分. 令

$$\Delta = E|\eta - (a\xi + b)|^2,$$

则

$$\begin{aligned} \Delta &= E|(\eta - E\eta) - a(\xi - E\xi)|^2 + |E\eta - aE\xi - b|^2 \\ &= D\eta + |a|^2 D\xi - \bar{a}\bar{b}_{\xi\eta} - ab_{\xi\eta} + |E\eta - aE\xi - b|^2 \\ &= D\eta - \frac{|b_{\xi\eta}|^2}{D\xi} + \left|a\sqrt{D\xi} - \frac{\overline{b_{\xi\eta}}}{\sqrt{D\xi}}\right|^2 + |E\eta - aE\xi - b|^2. \end{aligned}$$

故当

$$a = \frac{\overline{b_{\xi\eta}}}{D_\xi}, \qquad b = E\eta - aE\xi,$$

Δ 取得极小值时

$$\Delta = D_\eta - \frac{|b_{\xi\eta}|^2}{D\xi} = D\eta(1 - |r_{\xi\eta}|^2).$$

因此, $|r_{\xi\eta}| = 1$ 的充分必要条件是

$$\Delta = \int |\eta - (a\xi + b)|^2 dP = 0.$$

由积分的性质知

$$P(|\eta - (a\xi + b)| \neq 0) = 0.$$

此即所要证的结论. □

由 ii) 可知, $\Delta = D\eta(1 - |r_{\xi\eta}|^2)$ 表示用 $a\xi + b$ 作为 η 的近似值的精确程度, 也表示 ξ, η 之间线性关系的强弱. 而 Δ 又依赖于 $r_{\xi\eta}$. 因此 $r_{\xi\eta}$ 表示 ξ 与 η 之间线性关系强弱的程度.

定义 4 设 $\xi = (\xi_1, \cdots, \xi_n)$ 是概率场 $(\Omega, \mathscr{A}, P)$ 上的 n 维随机变量, 则称 $E\xi = (E\xi_1, \cdots, E\xi_n)$ 为 ξ 的**数学期望**, 称方阵

$$\begin{pmatrix} b_{11}, \cdots, b_{1n} \\ \cdots \\ b_{n1}, \cdots, b_{nn} \end{pmatrix}, b_{ij} = b_{\xi_i \xi_j}, i, j = 1, \cdots, n \tag{12}$$

为 ξ 的**相关方阵**.

性质 6 如果 $b_{ij} = b_{\xi_i \xi_j}, i, j = 1, \cdots n$ 有限, 则相关方阵非负定, 因而

$$\begin{vmatrix} b_{11}, \cdots, b_{1n} \\ \cdots \\ b_{n_1}, \cdots, b_{nn} \end{vmatrix} \geqslant 0.$$

证 对于任意一组复数 $t_1, \cdots, t_n$, 永远有

$$\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} t_i \bar{t}_j = \sum_{i=1}^{n}\sum_{j=1}^{n} E(\xi_i - E\xi_i)\overline{(\xi_j - E\xi_j)} t_i \bar{t}_j$$

$$= E\left|\sum_{i=1}^{n} t_i(\xi_i - E\xi_i)\right|^2 \geqslant 0.$$

由此可见, ξ 的相关方阵为非负定的, 因此, 行列式非负. □

如果将 $\xi=(\xi_1,\cdots,\xi_n)$ 了解为随机落入 n 维空间的点, 那么 $E\xi=(E\xi_1,\cdots,E\xi_n)$ 可以解释为 $(\xi_1,\cdots,\xi_n)$ 的平均位置. 至于 n 维随机变量的相关方阵, 则表达了各分量围绕它们的数学期望的疏散程度, 以及各分量间相关程度的一组数学特征的总和. 下面的性质可以说明这个解释 (也是性质 5 中 ii, 的推广).

以 B 表 n 维随机变量 $\xi=(\xi_1,\cdots,\xi_n)$ 的相关方阵, $r(B)$ 为 B 的秩, 则 $r(B)$ 也称为 **ξ的秩**. 于是有

性质 7　$(\xi_1,\cdots,\xi_n)$ 的秩为 $r(\leqslant n)$ 的充分必要条件是

i) 存在 $n-r$ 个线性无关的 n 维向量 $(t_1^{(l)},\cdots,t_n^{(l)})l=1,2,\cdots,n-r$, 使得

$$P\left(\sum_{k=1}^{n}t_k^{(l)}(\xi_k-E\xi_k)=0,l=1,\cdots,n-r\right)=1 \tag{13}$$

和

ii) 对于任何 $0\leqslant r_1<r$, 不存在 $n-r_1$ 个线性无关的 n 维向量具有性质 i).

这个性质说明 $(\xi_1,\cdots,\xi_n)$ 的秩为 r 的条件是 ξ 以概率 1 分布在一个 r 维超平面上, 而不可能以概率 1 分布在低于 r 维的超平面上. 这个性质我们不再证明, 读者可参考克拉美著的《统计学数学方法》(中译本)§22.5.

定义 5　设 ξ 是概率场 $(\Omega,\mathscr{A},P)$ 上的一个实随机变量, 我们规定下面的名称和记号:

$$\alpha_k=E\xi^k,k=0,1,2,\cdots \tag{14}$$

(如果存在) 称为 ξ 的 **k阶原点矩**;

$$\mu_k=E(\xi-E\xi)^k \qquad k=0,1,2,\cdots \tag{15}$$

(如果存在) 称为 ξ 的 **k阶中心矩**;

$$\beta_r=E|\xi|^r(r\geqslant 0) \tag{16}$$

称为 ξ 的 **r阶绝对矩**.

容易看出 ξ 的一阶原点矩就是 ξ 的数学期望,

$$\alpha_1=E\xi=m \tag{17}$$

当 ξ 是实随机变量时, 二阶中心矩就是它的方差, 即

$$\mu_2=E(\xi-E\xi)^2=D\xi. \tag{18}$$

中心矩与原点矩之间有简单的关系 (假定所涉及的矩都存在), 事实上, 由于

$$(\xi-E\xi)^n=\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}m^{n-k}\xi^k,$$

所以

$$\mu_n = E(\xi - E\xi)^n = \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}m^{n-k}\alpha_k. \tag{19}$$

而 $\alpha_0 = 1, \alpha_1 = m$, 得到

$$\mu_n = \sum_{k=2}^{n}(-1)^{n-k}\binom{n}{k}m^{n-k}\alpha_k + (-1)^{n-1}(n-1)m^n.$$

具体地写出来就是

$$\begin{aligned}
\mu_0 &= 1,\\
\mu_1 &= 0,\\
\mu_2 &= \alpha_2 - m^2,\\
\mu_3 &= \alpha_3 - 3m\alpha_2 + 2m^3,\\
\mu_4 &= \alpha_4 - 4m\alpha_3 + 6m^2\alpha_2 - 3m^4,\\
&\ \ \vdots
\end{aligned}$$

性质 8 如果 $E|\xi|^r < \infty$, 则对于 $r' \leqslant r, E|\xi|^{r'}$ 有限, 而且对于 $k \leqslant r, E\xi^k$ 存在并且有限.

证 因为当 $0 < r' < r$ 时

$$|\xi|^{r'} \leqslant 1 + |\xi|^r,$$

故有

$$E|\xi|^{r'} \leqslant 1 + E|\xi|^r.$$

这就证明了性质的前一部分, 而后一部分是前一部分的直接推论.

3.4.2 积分变换定理, 随机变量函数的分布律及数学期望的 L-S 积分表示

由于一般测度空间及一般的积分理论是比较广泛的, 掌握起来是比较困难的, 因此将概率论的问题归结到一类比较易于掌握的空间上来的想法是可取的.

因为分布律完全描述了随机变量的概率性质, 对于和 n 维随机变量有关的积分理论, 有可能变为在概率空间 $(R^{(n)}, \mathscr{B}^{(n)}, P_\xi)$ 上来讨论, 其中 P_ξ 是 ξ 的概率分布. 完成这个改变的基础是积分变换定理.

定义 6 若 f 是 $(\Omega, \mathscr{A})$ 到 $(X, \mathscr{B})$ 上的可测映射, μ 是 $(\Omega, \mathscr{A})$ 上的测度, 则称 μ_f

$$\mu_f(B) \triangleq \mu(f^{-1}(B)), B \in \mathscr{B} \tag{20}$$

为 μ 在 $(X, \mathscr{B})$ 上由 f **导出的测度**.

例 2　若 ξ 是概率场 $(\Omega,\mathscr{A},P)$ 上的 n 维实随机变量, 则 ξ 的概率分布 P_ξ 是 $(R^{(n)},\mathscr{B}^{(n)})$ 上由 ξ 导出的测度.

定理 1(积分变换定理)　设 f 是可测空间 $(\Omega,\mathscr{A})$ 到可测空间 $(X,\mathscr{B})$ 的可测变换, g 是 $(X,\mathscr{B})$ 上的可测函数, μ 是 $(\Omega,\mathscr{A})$ 上的测度, μ_f 是在 $(X,\mathscr{B})$ 上由 f 导出的测度, 则

$$\int_{f^{-1}(B)} g\circ f d\mu=\int_B g d\mu_f. \tag{21}$$

上式的意义是: 如等式一方的积分存在, 则另一方的积分也存在, 而且二者相等. 式中的 $g\circ f$ 是 g 和 f 的复合映射.

证　首先令 $g=\chi_F$, 其中 $F\in\mathscr{B}$, 则

$$\begin{aligned}\int_B \chi_F d\mu_f&=\mu_f(F\cap B)=\mu(f^{-1}(F)\cap f^{-1}(B))\\&=\int_{f^{-1}(B)}\chi_{f^{-1}(F)}d\mu=\int_{f^{-1}(B)}\chi_F\circ f d\mu.\end{aligned}$$

于是定理对于示性函数成立. 由积分的线性性质知定理对 g 是非负简单函数的情形成立. 若 g 是非负可测函数, 则存在简单函数序列 $\{g_n\},0\leqslant g_n\uparrow g$, 此时 $\{g_n\circ f\}$ 也是简单函数序列且 $0\leqslant g_n\circ f\uparrow g\circ f$, 因此定理对非负可测函数成立, 若 g 是实可测函数, 则

$$\int_{f^{-1}(B)} g^+\circ f d\mu=\int_B g^+ d\mu_f, \tag{22}$$

$$\int_{f^{-1}(B)} g^-\circ f d\mu=\int_B g^- d\mu_f, \tag{23}$$

而 $(g\circ f)^\pm=g^\pm\circ f$, $\int_B g d\mu_f$ 存在的充分必要条件是 (22), (23)式中有一有限, 它也是 $\int_{f^{-1}(B)} g^\circ f d\mu$ 存在的充分必要条件.

对于复可测函数 $g=g_1+ig_2(g_1,g_2$ 是实可测函数), 由于 $g\circ f=g_1\circ f+ig_2\circ f$, 应用实可测函数的结果, 可证定理仍然成立.　□

为了将随机变量函数的分布律及数学期望转化为 $(R^{(n)},\mathscr{B}^{(n)},P_\xi)$ 上的积分, 我们先给出 L–S 积分的定义.

定义 7　设 μ 是 $(R^{(n)},\mathscr{B}^{(n)})$ 上的 L–S 测度, f 是 $(R^{(n)},\mathscr{B}^{(n)})$(或 $(R^{(n)},\mathscr{B}_\mu^{(n)})$, $\mathscr{B}_\mu^{(n)}$ 是 $\mathscr{B}^{(n)}$ 对测度 μ 的完全化) 上的可测函数, 称 $(R^{(n)},\mathscr{B}_\mu^{(n)},\mu)$ 为 L–S空间, 若 F 是与 μ 对应的分布函数, 称 f 关于 μ(或关于 F) 的积分为 Lebesgue-Stieltjes **积分**, 简称 L–S 积分, 在 $R^{(n)}$ 上的积分记作

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1,\cdots,x_n)dF(x_1,\cdots,x_n)\text{或}\int_{R^{(n)}} f(x_1,\cdots,x_n)d\mu,$$

在集 $B \in \mathscr{B}^{(n)}$(或 $\mathscr{B}_\mu^{(n)}$) 上的积分记作

$$\int \underset{B}{\cdots} \int f(x_1, \cdots, x_n) dF(x_1, \cdots, x_n) \text{或} \int_B f(x_1, \cdots, x_n) d\mu,$$

在区间 $[a, b)$ 上的积分记作

$$\int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \cdots, x_n) dF(x_1, \cdots, x_n)$$

或

$$\int_{[a,b)} f(x_1, \cdots, x_n) d\mu,$$

其中 $a = (a_1, \cdots, a_n), b = (b_1, \cdots, b_n)$.

特别, 若 μ 为 n 维 Lebesgue 测度, 则称 $(R^{(n)}, \mathscr{B}_\mu^{(n)}, \mu)$ 为 Lebesgue **空间**, 简称L-**空间**, f 关于 μ 的积分称为 Lebesgue**积分**, 简称 L-**积分**. f 在 $R^{(n)}$, 集 B, 区间 $[a, b)$ 上的积分分别记作

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \cdots, x_n) dx_1 \cdots dx_n,$$

$$\int \underset{B}{\cdots} \int f(x_1, \cdots, x_n) dx_1 \cdots dx_n,$$

$$\int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \cdots, x_n) dx_1 \cdots dx_n.$$

现在我们来解决本节的中心问题，以下几个定理实质上是积分变换定理的直接推论。

定理 2 设 $\xi = (\xi_1, \cdots, \xi_n)$ 是 $(\Omega, \mathscr{A}, P)$ 上的 n 维实随机变量, 它的分布函数是 F, 则对于 n 维 Borel 集 G

$$P(\xi \in G) = \int \underset{G}{\cdots} \int dF(x_1, \cdots, x_n), \tag{24}$$

其中积分表示 L–S 积分.

证 在积分变换定理中令 $(X, \mathscr{B})$ 为 $(R^{(n)}, \mathscr{B}^{(n)}), f = \xi, g = 1, \mu = P$, 则 $\mu_f = P_\xi$ 因而

$$\begin{aligned} P(\xi \in G) &= P(\xi^{-1}(G)) \\ &= \int_{\xi^{-1}(G)} dP = \int_G dP_\xi = \int \underset{G}{\cdots} \int dF(x_1, \cdots, x_n). \end{aligned}$$ □

在 §2.3 定理 4 中, 我们已经看到随机变量的 Borel 可测函数的分布函数由随机变量的概率分布决定, 现在我们给出随机变量函数的分布函数的 L–S 积分表示.

定理 3　设 $\xi=(\xi_1,\cdots,\xi_n)$ 是 $(\Omega,\mathscr{A},P)$ 上的 n 维实随机变量, 它的分布函数是 $F(x_1,\cdots,x_n), g_k, k=1,\cdots,m$ 是 n 维实空间上的有限实 Borel 函数. 令 $\eta_k=g_k(\xi_1,\cdots,\xi_n), k=1,\cdots,m$, 则随机变量 $(\eta_1,\cdots,\eta_m)$ 的分布函数

$$F_{\eta_1\cdots\eta_m}(y_1,\cdots,y_m)=\int\cdots\int_G dF(x_1,\cdots,x_n), \tag{25}$$

其中积分区域

$$G=\{(x_1,\cdots,x_n):g_k(x_1,\cdots,x_n)<y_k,k=1,\cdots,m\}. \tag{26}$$

证　由下面事件的等式

$$\begin{aligned}\{\eta_1<y_1,\cdots,\eta_m<y_m\}=&\{\omega:\ g_1(\xi_1(\omega),\cdots,\xi_n(\omega))<y_1,\\&\cdots,g_m(\xi_1(\omega_1),\cdots,\xi_n(\omega))<y_m\}\\=&\{\omega:\ (\xi_1(\omega),\cdots,\xi_n(\omega))\in G\}=\{(\xi_1,\cdots,\xi_n)\in G\}.\end{aligned}$$

再由定理 2 有

$$\begin{aligned}&F_{\eta_1,\cdots,\eta_m}(y_1,\cdots,y_m)=P(\eta_1<y_1,\cdots,\eta_m<y_m)\\=&P((\xi_1,\cdots,\xi_n)\in G)=\int\cdots\int_G dF(x_1,\cdots,x_n).\end{aligned}$$

这就是定理所要证明的等式.　□

下面是随机变量函数的数学期望的 L–S 积分表示.

定理 4　设 $\xi=(\xi_1,\cdots,\xi_n)$ 是 $(\Omega,\mathscr{A},P)$ 上的 n 维实随机变量, 已知它的分布函数是 F, 又 g 是 n 维实空间上的有限 Borel 函数 (实值或复值), 则 $\eta=g(\xi_1,\cdots,\xi_n)$ 的数学期望存在的充要条件是 $g(x_1,\cdots,x_n)$ 关于 $F(x_1,\cdots,x_n)$ 的积分存在且

$$E\eta=Eg(\xi_1,\cdots,\xi_n)=\int\cdots\int g(x_1,\cdots,x_n)dF(x_1,\cdots,x_n). \tag{27}$$

证　由积分变换定理, 当下式两端的积分有一存在时,

$$\int_{\xi^{-1}(R^{(n)})}g\circ\xi dP=\int\cdots\int_{R^{(n)}}g(x_1,\cdots,x_n)dP_\xi. \tag{28}$$

而 (28) 式的左端为 $\int_\Omega g(\xi(\omega))dP=\int g(\xi_1,\cdots,\xi_n)dP=E\eta$. 由于 P_ξ 对应的分布

函数是 F, 故 (28) 式的右端就是

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1,\cdots,x_n)dF(x_1,\cdots,x_n).$$

关于 $E\eta$ 存在的充分必要条件也同时得出. □

推论 若 ξ 是 $(\Omega,\mathscr{A},P)$ 上的实随机变量, 其分布函数为 F, 则 ξ 的数学期望存在的充分必要条件是 $\int xdF(x)$ 存在, 且当 $\int xdF(x)$ 存在时,

$$E\xi=\int xdF(x). \tag{29}$$

在一些概率论的教科书中将 (29) 式作为数学期望的定义, 不过多半把 (29) 式右端的积分理解为 R–S 积分 (Riemann-Stieltjes 积分). (关于一元函数的 $R-S$ 积分的定义见 [17] 三卷一分册 548 目.)

习题及补充

1. 设 ξ,η 为实随机变量, ξ^2,η^2 的数学期望有限, 则 $D(\xi+\eta)=D\xi+D\eta$ 的充分与必要条件是 $E\xi\eta=E\xi\cdot E\eta$.

2. 随机变量 $(\xi_1,\cdots,\xi_n)$ 与 $(\eta_1,\cdots,\eta_n)$ 同分布, g 是 $R^{(n)}$ 上的 Borel 可测函数. 证明: 若 $Eg(\xi_1,\cdots,\xi_n)$ 存在, 则 $Eg(\eta_1,\cdots,\eta_n)$ 存在, 且

$$Eg(\xi_1,\cdots,\xi_n)=Eg(\eta_1,\cdots,\eta_n).$$

由此证明 $(\xi_1,\cdots,\xi_n)$ 与 $(\eta_1,\cdots,\eta_n)$ 有相同的数学期望、相关方阵及各阶矩.

3. 设 ξ 是一实随机变量, m 是一个实数, 满足以下条件:

$$P(\{\xi\leqslant m\})\geqslant\frac{1}{2},\quad P(\{\xi\geqslant m\})\geqslant\frac{1}{2}.$$

(称满足上述条件的 m 为 ξ 的中数.) 证明

i) 对于任意实数 a,

$$E|\xi-m|\leqslant E|\xi-a|.$$

ii) ξ 的中数、数学期望和方差之间有下述关系:

$$E\xi-\sqrt{D\xi}\leqslant m\leqslant E\xi+\sqrt{D\xi}.$$

4. 若 $\xi_1,\cdots,\xi_n$ 是 n 个随机变量, 它的分布函数分别为 $F_1(x),\cdots,F_n(x),\eta$ 是一随机变量, 分布函数是 $\frac{1}{n}(F_1(x)+\cdots+F_n(x))$, 设 r 是一个正数, 证明: 若

$E|\xi_k|^r, k=1,\cdots,n$ 有限, 则 $E|\eta|^r$ 有限, 且

$$E|\eta|^r = \frac{1}{n}\sum_{k=1}^{n} E|\xi_k|^r.$$

5. 设 ξ 是一个随机变量, 分布函数为 $F(x)$,

$$\xi_c = \begin{cases} \xi, & |\xi| < c, \\ c, & |\xi| \geqslant c. \end{cases}$$

试将 $E\xi_c, D\xi_c$ 用关于 $F(x)$ 的 L–S 积分表出.

6. 设 $F(x)$ 是一分布函数, 按定义

$$\int_a^b f(x)dF(x) = \int_{[a,b)} f(x)dF(x).$$

问 $\int_a^b f(x)dF(x)$ 与 $\int_{[a,b]} f(x)dF(x)$ 是否相等? 在什么情况下它们不相等?

7. 设随机变量 ξ 具有分布函数 $F(x)$, 求下列各随机变量的分布函数:

i) $\eta = a\xi + b, a, b$ 是实数;

ii) $\eta = \frac{1}{\xi}; (P(\{\xi = 0\}) = 0)$

iii) $\eta = \mathrm{t_g}\xi; \left(P\left(\left\{\xi = k\pi + \frac{\pi}{2}\right\}\right) = 0, k = \cdots, -1, 0, 1, \cdots\right)$

iv) $\eta = \cos\xi$.

8. 设 $(\xi_1, \cdots, \xi_n)$ 的分布函数为 $F(x_1, \cdots, x_n)$ 试求下述随机变量的分布函数.

i) $\eta = \max\{\xi_1, \cdots, \xi_n\}$;

ii) $\zeta = \min\{\xi_1, \cdots, \xi_n\}$.

再若 $\xi_1, \cdots, \xi_n$ 独立同分布, 试证 (η, ξ) 的分布函数 $F_{\eta\xi}(v, u)$ 是 $[F(v)]^n - g(u,v)[F(v) - F(u)]^n$, 其中 $F(v)$ 是 ξ_i 的分布函数, 而 $g(u,v) = \begin{cases} 1, & u < v. \\ 0, & u \geqslant v. \end{cases}$

9. 求证: 对于随机变量 ξ 及正值上升函数 $f(x), x \geqslant 0$, 恒有

$$P(|\xi| \geqslant x) \leqslant \frac{Ef(|\xi|)}{f(x)}, \quad (x > 0).$$

10. 设

$$F(x) = \begin{cases} 0, & x \leqslant 0, \\ \sigma^2, & 0 < x \leqslant 1, \\ \sigma^2 + \lambda, & x > 1, (\lambda > 0). \end{cases}$$

$f(x)$ 是 $R^{(1)}$ 上 Borel 可测函数, 试计算 L–S 积分

$$\int_{-\infty}^{\infty} f(x)dF(x).$$

§3.5 离散型和连续型随机变量

在 §2.3 我们粗略地叙述了离散型和连续型随机变量的定义, 在这一节将要重新给出它们的数学定义及数字特征的计算公式. 因为在理论上研究对连续型随机变量时常常使用 Lebesgue 积分, 但在实际计算时, 为了利用古典数学分析的已有结果, 则使用 Riemann 积分 (简称 R 积分), 因此在讨论连续型随机变量之前我们先对 L 积分与 R 积分进行比较, 弄清它们之间的关系.

3.5.1 离散型随机变量

定义 1 设 $\xi=(\xi_1,\cdots,\xi_n)$ 是概率场 $(\Omega,\mathscr{A},P)$ 上的 n 维随机变量, 且每一 ξ_k 只取有限个或可数个 (实或复) 数为值, 则称 ξ 为 n 维离散型随机变量.

我们回忆简单函数和初等函数的定义, 即知 n 维离散型随机变量就是 $(\Omega,\mathscr{A},P)$ 上的 n 个简单函数或初等函数所作成的取有限值的 n 维向量函数.

若 ξ_k 只取 $x_{k1},x_{k2},\cdots$ 为值, 则

$$\begin{aligned}P(\xi_1=x_{1l_1},\cdots,\xi_n=x_{nl_n})\ &=p_{l_1,\cdots,l_n},l_k=1,2,\cdots,\\ k\ &=1,\cdots,n,\end{aligned} \tag{1}$$

是 ξ 的分布阵列, 由它可以决定概率分布

$$P(\xi\in B)=\sum_{(x_{1l_1},\cdots,x_{nl_n})\in B} p_{l_1,\cdots,l_n}. \tag{2}$$

值得特别注意的是: 从实际问题中抽象出离散型随机变量并给出 (1) 式所标出的一切概率 $p_{l_1,\cdots,l_n}$ 时, 必须要求它满足:

$$p_{l_1,\cdots,l_n}\geqslant 0,\quad 对一切\quad l_k=1,2,\cdots,k=1,\cdots,n. \tag{3}$$

$$\sum_{l_1,\cdots,l_n} p_{l_1,\cdots,l_n}=1. \tag{4}$$

在第 2 章中我们讨论的二项分布、多项分布、Poisson 分布、多维 Poisson 分布的随机变量等都是离散型随机变量的例子. 今再举几例.

例 1 设有 N 个产品中, 其中有 M 个是废品, 今从中抽出 n 个, 令 ξ 表示抽出的 n 个产品中废品的个数, 则 ξ 为一只取 $0,1,\cdots,\min(n,M)$ 中的数为值的离散

型随机变量, 其分布列为

$$P(\xi=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},k=0,1,\cdots,\min(n,M). \tag{5}$$

这就是它的分布律, 称为超几何分布律.

例 2　若 ξ_r 表示某射手在射击试验中第 r 次射中的试验次数, 若每次射中的概率为 p, 则

$$P(\xi_r=r+k)=\binom{-r}{k}\cdot p^r(-q)^k, \tag{6}$$

其中

$$\binom{-r}{k}=\frac{(-r)(-r-1)\cdots(-r-k+1)}{k!},q=1-p.$$

ξ_r 的分布函数为

$$F_{\xi_r}(x)=\sum_{0\leqslant k<x-r}\binom{-r}{k}p^r(-q)^k.$$

这个函数称为负二项分布, 而当 $r=1$ 时称为几何分布 (这个分布的概率场的建立参看 §4.3 例 2).

现在我们给出离散型随机变量数字特征的计算公式.

定理 1　设 $\xi=(\xi_1,\cdots,\xi_n)$ 是 $(\Omega,\mathscr{A},P)$ 上的离散型 n 维随机变量, 所取的一切可能值为

$$(a_{1l_1},\cdots,a_{nl_n}),l_1,\cdots,l_n=1,2,\cdots,$$

且

$$P(\xi_1=a_{1l_1},\cdots,\xi_n=a_{nl_n})=p_{l_1,\cdots,l_n},l_1,\cdots,l_n=1,2,\cdots,$$

若 $f_1,\cdots,f_m$ 是 $R^{(n)}$ 上的有限值 Borel 函数, 则

$$\eta=(f_1(\xi_1,\cdots,\xi_n),\cdots,f_m(\xi_1,\cdots,\xi_n))$$

仍为离散型随机变量, 设它取的一切可能值为

$$(b_{1r_1},\cdots,b_{mr_m}),r_1,\cdots,r_m=1,2,\cdots,$$

则 η 的分布列为

$$\begin{aligned}P(f_1(\xi)=b_{1r_1},\cdots,f_m(\xi)=b_{mr_m})\\=\sum_{\substack{f_k(a_{1l_1},\cdots,a_{nl_n})=b_{kr_k}\\k=1,\cdots,m.}}p_{l_1,\cdots,l_n}.\end{aligned} \tag{7}$$

若 $g(x_1,\cdots,x_n)$ 是 $R^{(n)}$ 上的 Borel 函数 (实或复值), 则 $Eg(\xi_1,\cdots,\xi_n)$ 存在的充分必要条件是级数的和

$$\sum_{l_1,\cdots,l_n} g(a_{1l_1},\cdots,a_{nl_n})p_{l_1,\cdots,l_n} \tag{8}$$

存在, 且在 $Eg(\xi_1,\cdots,\xi_n)$ 存在时

$$Eg(\xi_1,\cdots,\xi_n)=\sum_{l_1,\cdots,l_n} g(a_{1l_1},\cdots,a_{nl_n})p_{l_1,\cdots,l_n}. \tag{9}$$

注 此处所说的级数的和存在, 比数学分析中的级数收敛要广一些. 是指正项的和与负项的和至少有一有限.

证 记 $(\xi_1,\cdots,\xi_n)$ 所取的一切可能值的集合为

$$\begin{aligned}B&=\{(a_{1l_1},\cdots,a_{nl_n}):l_1,\cdots,l_n=1,2,\cdots\}\\&=\sum_{l_1,\cdots,l_n}\{(a_{1l_1},\cdots,a_{nl_n})\}\epsilon\mathscr{B}^{(n)},\end{aligned}$$

其中 $\{(a_{1l_1},\cdots,a_{nl_n})\}$ 表示单点集. 那么

$$\begin{aligned}P_\xi(R^{(n)}-B)&=P(\{\omega:(\xi_1(\omega),\cdots,\xi_n(\omega))\epsilon R^{(n)}-B\})\\&=P(\varnothing)=0,\\P_\xi(\{(a_{1l_1},\cdots,a_{nl_n})\})&=P(\xi_1=a_{1l_1},\cdots,\xi_n=a_{nl_n})\\&=p_{l_1,\cdots,l_n}.\end{aligned}$$

于是

$$\begin{aligned}&P(f_1(\xi)=b_{1r_1},\cdots,f_m(\xi)=b_{mr_m})\\&\quad=\sum_{\substack{f_k(a_{1l_1},\cdots,a_{nl_n})=b_{kr_k}\\k=1,\cdots,m}}P(\xi_1=a_{1l_1},\cdots,\xi_n=a_{nl_n})\\&\quad=\sum_{\substack{f_k(a_{1l_1},\cdots,a_{nl_n})=b_{kr_k}\\k=1,\cdots,m}}p_{l_1,\cdots,l_n}.\end{aligned}$$

即 (7) 式成立. 利用数学期望的 L–S 积分表示知 $Eg(\xi_1,\cdots,\xi_n)$ 存在的充分必要条件是 $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(x_1,\cdots,x_n)dF(x_1,\cdots,x_n)$ 存在, 再利用不定积分的 σ-可加性

(§ 3, 推论 2), 当积分存在时,

$$\begin{aligned}Eg(\xi_1,\cdots,\xi_n) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1,\cdots,x_n)dF(x_1,\cdots,x_n)\\ &= \int_B g(x_1,\cdots,x_n)dP_\xi + \int_{R^{(n)}-B} g(x_1,\cdots,x_n)dP_\xi\\ &= \sum_{l_1,\cdots,l_n}\int_{\{(a_{1l_1},\cdots,a_{nl_n})\}} g(x_1,\cdots,x_n)dP_\xi\\ &= \sum_{l_1,\cdots,l_n} g(a_{1l_1},\cdots,a_{nl_n})pl_1,\cdots,l_n.\end{aligned}$$

即 $Eg(\xi_1,\cdots,\xi_n)$ 存在的充分必要条件是 (8) 中的级数和存在, 且存在时 (9) 成立.□

设 ξ 是取值 $a_1,a_2,\cdots,\eta$ 是取值 $b_1,b_2,\cdots$, 的离散型随机变量, 根据上面的定理, 立即得出

$$\begin{aligned}E\xi &= \sum_k a_k P(\xi=a_k),\\ D\xi &= \sum_k |a_k-E\xi|^2 P(\xi=a_k)\\ &= \sum_k |a_k|^2 P(\xi=a_k) - |E\xi|^2,\\ b_{\xi\eta} &= \sum_k\sum_l (a_k-E\xi)(\overline{b_2-E\eta})P(\xi=a_k,\eta=b_l),\end{aligned}$$

$f_\xi(t)=Ee^{it\xi}=\sum\limits_k e^{ita_k}P(\xi=a_k)$(当 ξ 为实随机变量时). 其中 $E\xi,D\xi$ 及 $f_\xi(t)$ 分别表示 ξ 的数学期望、方差及特征函数, $b_{\xi\eta}$ 表示 ξ 和 η 的相关矩.

通过计算不难求得在二项分布 $B(n,p)$ 中参数 np 即 $E\xi$, 而 $D\xi=np(1-p)$. 在 Poisson 分布中参数 λ 既是随机变量的数学期望又是方差. 对于这样一些分布, 这些参数完全决定了分布律.

3.5.2　L- 积分与 R- 积分的关系

关于 L-积分的定义在 § 4 定义 7 中已经给出, 现在我们先对 n 维空间上的 R 积分的有关概念作一些必要的说明 (参看 [17] 三卷一分册).

首先, R 积分是在可求积区域上定义的, 我们先从平面上的可求积区域及其面积同 L 集及其 L 测度作一对比. 在数学分析中, 先是定义了有限矩形 $[a,b)$ 的面积, 其中 $a=(a_1,a_2),b=(b_1,b_2)$:

$$S([a,b))=(b_1-a_1)(b_2-a_2),$$

然后定义有限个不相重叠的有限矩形 $I_1, \cdots, I_n$ 构成的图形 (不妨称之为折边矩形)$P = I_1 + \cdots + I_n$ 的面积

$$S(P) = \sum_{i=1}^{n} S(I_i).$$

对于任意一个有界区域 G, 我们考察包含在 G 的内部的折边矩形 P' 和包含 G 的折边矩形 P'', 如果

$$\sup_{P' \subset G} S(P') = \inf_{P'' \supset G} S(P''),$$

则称 G 为可求积区域, 并且它的面积定义为

$$S(G) = \sup_{P' \subset G} S(P') = \inf_{P'' \supset G} S(P'').$$

可以证明可求积区域都是 L 集, 并且它的面积就是它的 L 测度.

对于 n 维空间中的可求积区域, 情况与此完全类似, 只要把上面定义中的平面矩形改成 n 维空间中的长方体 $[a, b), a = (a_1, \cdots, a_n), b = (b_1, \cdots, b_n)$, 而它的体积定义为 $V([a, b)) = (b_1 - a_1) \cdots (b_n - a_n)$, 其他的叙述都适用于 n 维空间的情形. 这样一来, 可求积区域都是有界的 L 集, 而它的体积就是 L 测度. 因此以下我们把可求积区域 G 的体积就用 $\mu(G)$ 来表示, μ 是 L 测度.

可求积区域 G 上函数 $f(x_1, \cdots, x_n)$ 的 R 积分 (n 重积分) 定义如下: 将 G 划分成有限个互不相交的可求积区域 $G_1, \cdots, G_m$ 的和集, 将此分划记为 T

$$G = G_1 + \cdots + G_m$$

$G_i, i = 1, \cdots, m$ 的直径中的最大者记作 $l(T)$, 在每一 G_i 中任取一点 $(x_{i_1}, \cdots, x_{i_n})$ ϵG_i, 如果

$$I = \lim_{l(T) \to 0} \sum_{i=1}^{m} f(x_{i_1}, \cdots, x_{i_n}) \mu(G_i) \tag{10}$$

存在 (有限), 且与划分方式及 $(x_{i_1}, \cdots, x_{i_n})$ 的选择无关, 则称 $f(x_1, \cdots, x_n)$在 G 上 R 可积, 并称 I 为 $f(x_1, \cdots, x_n)$ 在 G 上的积分, 记作

$$\int \cdots \int_G f(x_1, \cdots, x_n) dx_1 \cdots dx_n = I. \tag{11}$$

可以证明 $f(x_1, \cdots, x_n)$ 在 G 上 R 可积的充分与必要条件是

$$\lim_{l(T) \to 0} \sum_{i=1}^{m} M_i \mu(G_i) = \lim_{l(T) \to 0} \sum_{i=1}^{m} m_i \mu(G_i), \tag{12}$$

其中

$$M_i = \sup_{(x_1,\cdots,x_n)\epsilon G_i} f(x_1,\cdots,x_n),$$

$$m_i = \inf_{(x_1,\cdots,x_n)\epsilon G_i} f(x_1,\cdots,fx_n). \tag{13}$$

现在我们比较可求积区域上的 L 积分与 R 积分.

定理 2　如果可测函数 f 在可求积区域 G 上 R 可积, 则 f 在 G 上也是 L 可积的, 而且 f 在 G 上两种意义下的积分相等, 即

$$\int\cdots\int_G f(x_1,\cdots,x_n)dx_1\cdots dx_n = \int_G f(x_1,\cdots,x_n)d\mu, \tag{14}$$

其中 μ 是 L 测度.

注　在这一小节对 R 积分与 L 积分在表示法上有所区别, (14) 式左端表示 R 积分, 右端表示 L 积分.

证　由于 f 在 G 上 R 可积, 因此 f 在 G 上有界, 同时 $\mu(G)$ 有限, 所以 f 在 G 上 L 可积, 于是只要证明 (14) 式成立.

对于 G 的每一分划 $T, G = G_1 + \cdots + G_m$, 令

$$\underline{f}_T(x_1,\cdots,x_n) = \sum_{i=1}^{m} m_i\chi_{G_i}(x_1,\cdots,x_n),$$

$$\overline{f}_T(x_1,\cdots,x_n) = \sum_{i=1}^{M} m_i\chi_{G_i}(x_1,\cdots,x_n),$$

其中 m_i, M_i 的意义如 (13), 显然 $\underline{f}_T$ 和 $\overline{f}_T$ 是简单函数且在 G 上

$$\underline{f}_T \leqslant f \leqslant \overline{f}_T,$$

于是

$$\sum_{i=1}^{m} m_i\mu(G_i) = \int_G \underline{f}_T d\mu \leqslant \int_G f d\mu \leqslant \int_G \overline{f}_T d\mu = \sum_{i=1}^{m} M_i\mu(G_i).$$

令 $l(T) \to 0$ 由 (12) 即得 (14).　□

定理 2 对于无界区域也成立. 在无界区域上的 R 积分 (n 重广义积分) 是如下定义的:

设 G 是一无界区域, 用任一单连通的可求积区域 G' 截割它, $G \cap G'$ 是可求积区域. 若 $f(x_1,\cdots,x_n)$ 在每一 $G \cap G'$ 上可积, 且当 G' 的边界带着它所有的点移到无穷远去; 使自原点到这边界上的点的最小距离 R 增大到无穷, $G \cap G'$ 逐渐笼罩住

区域 G 的所有点时, $\lim\limits_{R\to\infty}\int_{G\cap G'}\cdots\int f(x_1,\cdots,x_n)dx_1,\cdots,dx_n$ 存在且有限, 则称它为无界区域 G 上函数 $f(x_1,\cdots,x_n)$ 的 (广义) 积分. 记作

$$\int\cdots\int_G f(x_1,\cdots,x_n)dx_1\cdots dx$$

可以证明, 当 $n\geqslant 2$ 时 $f(x_1,\cdots,x_n)$ 在无界区域 G 上 R 可积当且仅当 $|f(x_1,\cdots,x_n)|$ 在 G 上 R 可积 (即在 G 上 R 绝对可积). (参看 [17] 第三卷第一分册 587, 588 目)

定理 3 如果可测函数 $f(x_1,\cdots,x_n)$ 在无界区域 G 上 R 绝对可积, 且在 G 的每一有界可求积部分 R 可积, 则 $f(x_1,\cdots,x_n)$ 在 G 上 L 可积, 且在 G 上 $f(x_1,\cdots,x_n)$ 的两种意义的积分相等, 即

$$\int\cdots\int_G f(x_1,\cdots,x_n)dx_1\cdots dx_n=\int_G f(x_1,\cdots,x_n)d\mu.$$

证 取 $G_k=[(-k,\cdots,-k),(k,\cdots,k))$, 由于 $|f(x_1,\cdots,x_n)|$ 在 G 上 R 可积, 故

$$\begin{aligned}&\int\cdots\int_G |f(x_1,\cdots,x_n)|dx_1\cdots dx_n\\=&\lim_{k\to\infty}\int\cdots\int_{G\cap G_k}|f(x_1,\cdots,x_n)|dx_1\cdots dx_n\end{aligned}\tag{15}$$

有限. 由定理 2

$$\begin{aligned}&\int\cdots\int_{G\cap G_k}|f(x_1,\cdots,x_n)|dx_1\cdots dx_n\\=&\int_{G\cap G_k}\cdots|f(x_1,\cdots,x_n)|d\mu,\end{aligned}\tag{16}$$

而在 G 上, $0\leqslant|f(x_1,\cdots,x_n)|\chi_{G_k}(x_1,\cdots,x_n)\uparrow|f(x_1,\cdots,x_n)|,(k\to\infty)$, 利用单调收敛定理及 (16), (15) 式即得

$$\begin{aligned}&\int_G|f(x_1,\cdots,x_n)|d\mu\\=&\lim_{k\to\infty}\int_G|f(x_1,\cdots,x_n)|\chi_{G_k}(x_1,\cdots,x_n)d\mu\\=&\lim_{k\to\infty}\int_{G\cap G_k}|f(x_1,\cdots,x_n)|d\mu\end{aligned}$$

$$= \lim_{k\to\infty} \underset{G\cap G_k}{\int\cdots\int} |f(x_1,\cdots,x_n)|dx_1\cdots,ax_n$$

$$= \underset{G}{\int\cdots\int} |f(x_1,\cdots,x_n)|dx_1\cdots dx_n < \infty.$$

故 $|f(x_1,\cdots,x_n)|$ 在 G 上 L 可积, 因而 $f(x_1,\cdots,x_n)$ 在 G 上 L 可积.

其次, 由于对一切 $(x_1,\cdots,x_n)\in G$,

$$|f(x_1,\cdots,x_n)\chi_{G_k}(x_1,\cdots,x_n)| \leqslant |f(x_1,\cdots,x_n)|,$$

又 f 在 G 上 L 可积, 而且在 G 上

$$f(x_1,\cdots,x_n)\chi_{G_k}(x_1,\cdots,x_n) \to f(x_1,\cdots,x_n), (k\to\infty)$$

利用控制收敛定理及定理 2 , 就有

$$\int_G f(x_1,\cdots,x_n)d\mu$$

$$= \lim_{k\to\infty}\int_G f(x_1,\cdots,x_n)\chi_{G_k}(x_1,\cdots,x_n)d\mu$$

$$= \lim_{k\to\infty}\int_{G\cap G_k} f(x_1,\cdots,x_n)d\mu$$

$$= \lim_{k\to\infty} \underset{G\cap G_k}{\int\cdots\int} f(x_1,\cdots,x_n)dx_1\cdots dx_n$$

$$= \underset{G}{\int\cdots\int} f(x_1,\cdots,x_n)dx_1\cdots dx_n.$$ □

注　上面所给的广义 n 重积分在 $n=1$ 时和 $n>1$ 时具有不同的性质. 一元函数关于无界区域的广义积分, 绝对可积蕴含可积, 但可积并不能保证绝对可积, 而当 $n\geqslant 2$ 时可积与绝对可积等价. 在定理 3 中为了保证对 $n=1$ 也成立, 所以我们假设 $|f(x_1,\cdots,x_n)|$ 可积, 其实对于 $n\geqslant 2$ 来说, 只假定 $f(x_1,\cdots,x_n)$ 可积就够了.

下面给出一个可积但不绝对可积的例子.

例 3　$f(x)=\dfrac{\sin x}{x}$ 在 $[0,\infty)$ 上 R 可积, 但不绝对可积, 且在 L 积分意义下也不可积 (积分不存在). 事实上,

$$\int_0^\infty \frac{\sin x}{x}dx = \lim_{b\to\infty}\int_0^b \frac{\sin x}{x}dx = \frac{\pi}{2},$$

而

$$\int_0^\infty \frac{|\sin x|}{x} dx = \lim_{b\to\infty} \int_0^b \frac{|\sin x|}{x} dx = +\infty.$$

(参看 [17] 第二卷第三分册 439, 455 目.) 这时不能推出 $\frac{\sin x}{x}$ 在 $[0,\infty)$ 上 L 可积. 因为如果 $\frac{\sin x}{x}$ 在 $[0,\infty)$ 上 L 可积, 利用控制收敛定理

$$\int_{[0,\infty)} \frac{|\sin x|}{x} d\mu = \lim_{b\to\infty} \int_0^b \frac{|\sin x|}{x} d\mu,$$

在有限区间上

$$\int_{[0,b)} \frac{|\sin x|}{x} d\mu = \int_0^b \frac{|\sin x|}{x} dx.$$

因此得到

$$\lim_{b\to\infty} \int_0^b \frac{|\sin x|}{x} dx = \int_{[0,\infty)} \frac{|\sin x|}{x} d\mu < \infty.$$

这与 $\left|\frac{\sin x}{x}\right|$ 在 $[0,\infty)$ 上的 R 积分不存在 $(=+\infty)$ 矛盾. □

上面的推理完全适用于一般情况. 如果一元函数 $f(x)$ 在无穷区间上 R 可积, 但非绝对 R 可积, 则 $f(x)$ 非 L 可积, 并且可以证明 L 积分不存在.

对于无界函数的 R 积分 (广义 n 重积分) 也有类似于定理 3 的结果, 不再在此详述了.

为了应用的方便, 我们给出以下

推论 1 如果在有限矩形 $[(a_1,\cdots,a_n),(b_1,\cdots,b_n)]$ 上 $f(x_1,\cdots,x_n)$ 连续或在有限个分片光滑的 $n-1$ 维曲面上有不连续点 (但保持有界), 则 $f(x_1,\cdots,x_n)$ 在 L 积分及 R 积分两种意义下均可积且两种积分相等.

推论 2 如果在 $R^{(n)}$ 上 $f(x_1,\cdots,x_n)$ 连续, 或在任意有限区间上不连续点分布在有限个分片光滑的 $n-1$ 维曲面上 (但保持有界), 且 $f(x_1,\cdots,x_n)$ 在 $R^{(n)}$ 上 R 绝对可积, 则 f 在 $R^{(n)}$ 上 L 可积, 且 f 在 $R^{(n)}$ 上两种意义的积分相等.

下面我们对 L 积分给出类似于 R 积分的变量替换公式.

设 $(R_x^{(n)},\mathscr{B}_x^{(n)},\Lambda_x)$ 和 $(R_t^{(n)},\mathscr{B}_t^{(n)},\Lambda_t)$ 都是 Lebesgue 测度空间, 设 $D_x\in\mathscr{B}_x^{(n)}$, $D_t\in\mathscr{B}_t^{(n)}$, 在 D_x 和 D_t 之间存在一个一一对应:

$$\begin{aligned} x_1 &= x_1(t_1,\cdots,t_n), \\ &\cdots \qquad\qquad (t_1,\cdots,t_n)\in D_t, \\ x_n &= x_n(t_1,\cdots,t_n), \end{aligned}$$

函数 $x_i(t_1,\cdots,t_n), i=1,\cdots,n$, 在区域 D_i 上存在连续偏导数, 假定函数行列式

$$J = \frac{D(x_1,\cdots,x_n)}{D(t_1,\cdots,t_n)} \neq 0, (t_1,\cdots,t_n)\in D_t.$$

再设 D_x 和 D_t 是由光滑或逐片光滑的曲面围起来的, 或者是整个 n 维空间或半空间. 对于任意有限区间 $[a,b)$, 令

$$f(x_1,\cdots,x_n)=\chi_{[a,b)\cap D_x}(x_1,\cdots,x_n),(x_1,\cdots,x_n)\in D_x,$$

由 f 有界, 其不连续点分布在有限个逐片光滑的 $n-1$ 维曲面上. 根据 [17] 第三卷第二分册 649 目的结果 (对变量替换公式中的函数 f 推广至在有限个逐片光滑的 $n-1$ 维曲面上有不连续点的有界函数) 有

$$\begin{aligned}&\int\underset{D_x}{\cdots}\int f(x_1,\cdots,x_n)dx_1\cdots dx_n\\=&\int\underset{x^{-1}(D_x)}{\cdots}\int f(x_1(t)\cdots,x_n(t))|J|dt_1\cdots dt_n.\end{aligned}$$

上述积分既可理解为 R 积分, 又可理解为 L 积分. 因而上式即

$$\int_{D_x\cap[a,b)}d\Lambda_x=\int_{x^{-1}(D_x)}\chi_{[a,b)\cap D_x}(x_1(t),\cdots,x_n(t))|J(t)|d\Lambda_t.$$

即对一切 $a,b\in R^{(n)}$,

$$\begin{aligned}\Lambda_x(D_x\cap[a,b))&=\int_{x^{-1}(D_x)}\chi_{x^{-1}([a,b)\cap D_x)}(t)|J(t)|d\Lambda_t\\&=\int_{x^{-1}(D_x\cap[a,b))}|J(t)|d\Lambda_t.\end{aligned}$$

通过取极限, 上式对 $a=-\infty,b=+\infty$ 的情形均成立, 因而

$$\Lambda_x(D_x\cap B)=\int_{x^{-1}(D_x\cap B)}|J(t)|d\Lambda_t,B\in\mathbb{C}^{(n)}$$

成立. 由于逆象的性质、不定积分的 σ-可加性以及 Λ_x 是 $D_x\cap\mathscr{B}_x^{(n)}$ 上的测度知上式两端都是定义在 $D_x\cap\mathbb{C}^{(n)}$ 上的测度, 因而可以扩张为 $D_x\cap\mathscr{B}_x^{(n)}$ 上的测度, 且扩张是唯一的, 即对一切 $B\in D_x\cap\mathscr{B}_x^{(n)}$ 有

$$\Lambda_x(B)=\int_{x^{-1}(B)}|J(t)|d\Lambda_t.$$

于是对 $(D_x,D_x\cap\mathscr{B}_x^{(n)})$ 上非负简单函数 $f=\sum\limits_{i=1}^{n}a_i\chi_{A_i}$ 有

$$\int_{D_x}f(x)d\Lambda_x=\int_{D_x}\sum_{i=1}^{m}a_i\chi_{A_i}(x)d\Lambda_x$$

$$
\begin{aligned}
&=\sum_{i=1}^{m} a_i \Lambda_x(D_x\cap A_i)=\sum_{t=1}^{m} a_i\int_{x^{-1}(D_x\cap A_i)}|J(t)|d\Lambda_t\\
&=\int_{x^{-1}(D_x)}\sum a_i\chi_{x^{-1}(A_i)}(t)|J(t)|d\Lambda_t\\
&=\int_{x^{-1}(D_x)}\sum_{t=1}^{m} a_i\chi_{A_i}(x(t))|J(t)|d\Lambda_t.\\
&=\int_{x^{-1}(D_x)} f(x(t))|J(t)|d\Lambda_t.
\end{aligned}
$$

使用单调收敛定理知对一切非负可测函数

$$
\int_{D_x} f(x)d\Lambda_x=\int_{x^{-1}(D_x)} f(x(t))|J(t)|d\Lambda_t \tag{17}
$$

成立, 从而对一切积分存在的实、复可测函数上述变量替换公式成立.

3.5.3 连续型随机变量

定义 2 设 $\xi=(\xi_1,\cdots,\xi_n)$ 是概率场 $(\Omega,\mathscr{A},P)$ 上的 n 维实随机变量, 它的分布函数是 $F(x_1,\cdots,x_n)$, 若有一 $(R^{(n)},\mathscr{B}^{(n)})$ 上的非负可测函数 $p(x_1,\cdots,x_n)$ 存在, 使得对任何 $x=(x_1,\cdots,x_n)\in R^{(n)}$, 有

$$
F(x_1,\cdots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} p(t_1,\cdots,t_n)dt_1\cdots dt_n \tag{18}
$$

成立, 则称 $\xi=(\xi_1,\cdots,\xi_n)$ 是**连续型随机变量**, 而 $p(x_1,\cdots,x_n)$ 称为 ξ 的**概率分布密度**(简称分布密度). (18) 式右端的积分为 L 积分, 当 $p(x_1,\cdots,x_n)$ 的 R 积分存在时也是 R 积分.

注 由 §5.2 的结果, 今后 $R^{(n)}$ 上的 L 积分与 R 积分在积分符号上不加区别.

关于连续型随机变量, 有很多在实际应用中重要的例子, 我们在第 2 章 §2.1 中所举的均匀分布、正态分布及 n 维正态分布便是其中最重要的几例. 其他的例子请参看 [3].

现在我们先给出连续型随机变量概率分布的 L 积分表示以及连续型随机变量函数的分布律及数学期望的 L 积分表示式.

定理 4 若 $\xi=(\xi_1,\cdots,\xi_n)$ 为具有分布密度的连续型 n 维随机变量, 则

$$
P(\xi\in G)=P_\xi(G)=\int\underset{G}{\cdots}\int p(t_1,\cdots,t_n)dt_1\cdots dt_n, \tag{19}
$$

又若 $g_1(x_1,\cdots,x_n),\cdots,g_m(x_1,\cdots,x_n)$ 是 m 个实 Borel 可测函数, 而 $\eta_k=g_k(\xi_1,\cdots,\xi_n),k=1,\cdots,m$, 则 $(\eta_1,\cdots,\eta_m)$ 的分布函数

$$F_{\eta_1,\cdots,\eta_m}(y_1,\cdots,y_m)=\underset{\substack{g_k(x_1,\cdots,x_n)<y_k\\k=1,\cdots,m.}}{\int\cdots\int}p(x_1,\cdots,x_n)dx_1\cdots dx_n. \qquad (20)$$

上述积分为 L 积分, 当 $p(x_1,\cdots,x_n)$ R 可积且积分区域为可求积区域时可以理解为 R 积分.

证　根据第 2 章 §2.3 知 P_ξ 是 $\mathscr{B}^{(n)}$ 上的概率, 因而是 σ-有限测度, 而由不定积分的 σ-可加性 (§2.3 定理 2 推论 2) 知

$$\varphi(A)=\int\underset{A}{\cdots}\int p(t_1,\cdots,t_n)dt_1\cdots dt_n, A\in\mathscr{B}^{(n)}$$

也是一个 $\mathscr{B}^{(n)}$ 上的 σ-有限测度, 并且对任何 $a,b\in R^{(n)},a\leqslant b$.

$$\begin{aligned}P_\xi([a,b))=\Delta_{b,a}F&=\int\underset{[a,b)}{\cdots}\int p(x_1,\cdots,x_n)dx_1\cdots dx_n\\&=\varphi([ab)),\end{aligned}$$

其中 F 是 ξ 的分布函数, 第二个等式可由积分的性质推得, 再由第 2 章 §2.3 分布函数唯一决定 L–S 测度知

$$P_\xi(A)=\varphi(A), A\in\mathscr{B}^{(n)}.$$

(19) 式即获证明. 而

$$\begin{aligned}P(\eta_1<y_1,\cdots,\eta_n<y_n)&=P(g_k(\xi_1,\cdots,\xi_n)<y_k,k=1,\cdots,m)\\&=P((\xi_1,\cdots,\xi_n)\in G)=P_\xi(G),\end{aligned}$$

其中 $G=\{(x_1,\cdots,x_n):g_k(x_1,\cdots,x_n)<y_k,k=1,\cdots,m\}$, 因而由 (19) 推得 (20). □

下面我们推导一个求连续型随机变量函数的分布密度公式, 它在实际计算中是很有用的.

定理 5　设 $\xi=(\xi_1,\cdots,\xi_n)$ 是连续型随机变量, 其分布密度为 $p_\xi(x_1,\cdots,x_n)$, $R^{(n)}$ 上的 n 元可测函数 $y=(y_1,\cdots,y_n)$, 其中 $y_i(x_1,\cdots,x_n),i=1,\cdots,n$, 满足下述条件: 除 $R^{(n)}$ 中的 L-零测集 N 而外, 存在至多可数个两两不交的可求积区域 $\{D_{xk}\}_{k\geqslant1}$, 使 $R^{(n)}-N=\sum_{k=1}^{\infty}D_{xk}$, 对一切 k, 函数 y 把 D_{xk} 中的点一对一地映射到 $R_y^{(n)}$ 中的可求积区域 D_{yk} 上 (不同的 k 未必不相交), 在每一 D_{xk} 上, 设 y的逆映射为 $x^{(k)}=\left(x_1^{(k)},\cdots,x_n^{(k)}\right)$, 其中 $x_i^{(k)}=x_i^{(k)}(y_1,\cdots,y_n),i=1,\cdots,n$, 是 D_{yk} 上的连续函数, 且具有连续的偏导数, 函数行列式

$$J_k(y_1,\cdots,y_n)=\frac{D(x_1,\cdots,x_n)}{D(y_1,\cdots,y_n)}\neq 0.$$

若 $\eta=(\eta_1,\cdots,\eta_n)$, 其中 $\eta_i=y_i(\xi_1,\cdots,\xi_n), i=1,\cdots,n$, 则 η 是连续型随机变量, 其分布密度为

$$p_\eta(y_1,\cdots,y_n)=\sum_{k=1}^{\infty}\chi_{D_{yk}}(y_1,\cdots,y_n)p_\xi(x_1^{(k)}(y_1,\cdots,y_n),\\ \cdots,x_n^{(k)}(y_1,\cdots,y_n))|J_k(y_1,\cdots,y_n)|. \tag{21}$$

证 由定理 4 不难求得, 对任意 $B\in\mathscr{B}_y^{(n)}$

$$P(\eta\in B)=\int\underset{G}{\cdots}\int p_\xi(x_1,\cdots,x_n)dx_1\cdots dx_n,$$

其中 $G=\{(x_1,\cdots,x_n):(y_1(x_1,\cdots,x_n),\cdots,y_n(x_1,\cdots,x_n))\in B\}\in\mathscr{B}^{(n)}$. 于是

$$P(\eta\in B)=\sum_{k=1}^{\infty}\int\underset{G\cap D_{xk}}{\cdots}\int p_\xi(x_1,\cdots,x_n)dx_1\cdots dx_n.$$

利用 §5.2 导出的 L 积分变量替换公式, 以及 $x^{-1}(G\cap D_{xk})=y(G\cap D_{xk})=B\cap D_{yk}$ 有

$$\begin{aligned}P(\eta\in B)=&\sum_{k=1}^{\infty}\int\underset{B\cap D_{yk}}{\cdots}\int p_\xi(x_1(y_1,\cdots,y_n),\cdots,\\&x_n(y_1,\cdots,y_n))|J_k(y_1,\cdots,y_n)|dy_1\cdots dy_n\\=&\int_B\cdots\int\sum_{k=1}^{\infty}\chi_{D_{yk}}(y_1,\cdots,y_n)|J_k(y_1,\cdots,\\&y_n)|p_\xi(x_1^{(k)}(y_1,\cdots,y_n),\cdots,x_n^{(k)}(y_1,\cdots,y_n)|dy_1\cdots dy_n.\end{aligned}$$

由于上式对一切 $B\in\mathscr{B}_y^{(n)}$ 成立, 且被积函数为非负 Lebesgue 可测函数, 因而 η 为连续型随机变量且

$$p_\eta(y_1,\cdots,y_n)=\sum_{k=1}^{\infty}\chi_{D_{yk}}(y_1,\cdots,y_n)|J_k(y_1,\cdots,y_n)|\\ \cdot p_\xi(x_1^{(k)}(y_1,\cdots,y_n),\cdots,x_n^{(k)}(x_1,\cdots,x_n)).$$

□

例 4 (连续型随机变量的线性变换). 设 $\xi=(\xi_1,\cdots,\xi_n)$ 是具有分布密度 $p(x_1,\cdots,x_n)$ 的随机变量

$$\eta_k = \sum_{j=1}^{n} c_{kj}\xi_j + b_k, k = 1, \cdots, n, \tag{22}$$

且 $n \times n$ 方阵 $C = (c_{kj})$ 的行列式不为零, 则 $\eta = (\eta_1, \cdots, \eta_n)$ 为连续型随机变量, 其分布密度

$$p_\eta(y_1, \cdots, y_n) = \frac{1}{||C||} p\left(\sum_{k=1}^{n} c'_{1k}(y_k - b_k), \cdots, \sum_{k=1}^{n} c'_{nk}(y_k - b_k) \right), \tag{23}$$

其中 $(c'_{kl}) = C^{-1}$, $||C||$ 表示行列式 $|C|$ 的绝对值.

事实上, 作变量代换

$$y_k = \sum_{l=1}^{n} c_{kl}x_l + b_k, k = 1, \cdots, n,$$

则

$$x_l = \sum_{k=1}^{n} c'_{lk}(y_k - b_k), l = 1, \cdots, n.$$

这一变换的函数行列式

$$\frac{D(x_1, \cdots, x_n)}{D(y_1, \cdots, y_n)} = |C^{-1}| = \frac{1}{|C|}.$$

在 $R^{(n)}$ 上不为零, 且变换是一对一的, 根据定理 5 即得所求.

例 5　设 (ξ_1, ξ_2) 的分布密度为 $p(x_1, x_2)$, 则 $\xi_1 \cdot \xi_2$ 是连续型随机变量且其分布密度

$$p_{\xi_1\xi_2}(x) = \int_{-\infty}^{\infty} \frac{1}{|z|} p\left(\frac{x}{z}, z\right) dz.$$

事实上, 设 $y_1(x_1, x_2) = x_1 \cdot x_2, y_2(x_1, x_2) = x_2, N = \{(x_1, x_2) : x_2 = 0, (x_1, x_2) \in R^{(2)}\}, D_{x1} = \{(x_1, x_2) : x_2 < 0\}, D_{x2} = \{(x_1, x_2) : x_2 > 0\}$, 在 $D_{xk}, k = 1, 2$ 上, 映射 y 分别将 D_{x1} 映成 $D_{y1} = \{(y_1, y_2) : y_2 < 0\}$, 将 D_{x2} 映成 $D_{y2} = \{(y_1, y_2) : y_2 > 0\}$, 而逆映射

$$x_1(y_1, y_2) = \frac{y_1}{y_2},$$
$$x_2(y_1, y_2) = y_2.$$

函数行列式

$$J_k(y_1, y_2) = \frac{1}{y_2}, \quad k = 1, 2.$$

在 D_{y1}, D_{y2} 上不为零, 连续, 因而根据定理 5, 设 $\eta_1 = \xi_1\xi_2, \eta_2 = \xi_2$, 则

$$p_{\eta_1,\eta_2}(y_1, y_2) = \frac{1}{|y_2|} p\left(\frac{y_1}{y_2}, y_2\right),$$

而 $\eta_1=\xi_1\cdot\xi_2$ 的分布密度为

$$p_{\xi_1\cdot\xi_2}(x)=\int_{-\infty}^{\infty}\frac{1}{|z|}p\left(\frac{x}{z},z\right)dz.$$

例 5 为我们提供了一个方法, 当 $(\eta_1,\cdots,\eta_m), m<n$ 时我们仍然可以利用定理 5 计算 $(\eta_1,\cdots,\eta_m)$ 的分布密度, 只要适当地添加 $n-m$ 个 $(\xi_1,\cdots,\xi_n)$ 的连续函数即可.

进一步的例子, 读者可参看文献 [17] 中的 § 2.8.

定理 6 设 $\xi=(\xi_1,\cdots,\xi_n)$ 是具有分布密度 $p(x_1,\cdots,x_n)$ 的连续型随机变量, $g(x_1,\cdots,x_n)$ 是 n 元 Borel 可测函数, 则 $Eg(\xi_1,\cdots,\xi_n)$ 存在的充分必要条件是 $g(x_1,\cdots,x_n)p(x_1,\cdots,x_n)$ 在 $R^{(n)}$ 上的 L 积分存在, 且在 $Eg(\xi_1,\cdots,\xi_n)$ 存在时

$$Eg(\xi_1,\cdots,\xi_n)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(x_1,\cdots,x_n)p(x_1,\cdots,x_n)dx_1\cdots dx_n. \tag{24}$$

这一定理是下面关于一般积分性质的直接推论.

定理 7 设 μ 是可测空间 $(\Omega,\mathscr{A})$ 上的测度, λ 是 $(\Omega,\mathscr{A})$ 上某一非负可测函数 $p(\omega)$ 的不定积分 (因而是 $\mathscr{A}$ 上的一个测度):

$$\lambda(A)=\int_A p(\omega)d\mu(\omega), A\in\mathscr{A}.$$

则 $(\Omega,\mathscr{A})$ 上可测函数 $g(\omega)$ 的积分 $\int g(\omega)d\lambda(\omega)$ 存在当且仅当 $\int g(\omega)p(\omega)d\mu$ 存在, 且在积分存在时, 对任意 $A\in\mathscr{A}$,

$$\int_A g(\omega)d\lambda(\omega)=\int_A g(\omega)p(\omega)d\mu(\omega). \tag{25}$$

证 首先当 g 是示性函数时, 即 $g=\chi_B, B\in\mathscr{A}$, (25) 式成立, 即

$$\begin{aligned}\int_A g(\omega)d\lambda(\omega)&=\int_A\chi_B(\omega)d\lambda(\omega)=\lambda(A\cap B)\\&=\int_{A\cap B}p(\omega)d\mu(\omega)=\int_A\chi_B(\omega)p(\omega)d\mu(\omega)\\&=\int_A g(\omega)p(\omega)d\mu(\omega).\end{aligned}$$

从而由积分的线性性质及单调收敛定理知 (25) 式对一切非负可测函数成立.

设 g 为实可测函数, g^+,g^- 分别为其正部和负部, 则 g^+p,g^-p 分别为 gp 的正部和负部. 由已证结论易见 $\int gd\lambda$ 存在的充分必要条件是 $\int g^+d\lambda, \int g^-d\lambda$ 不同时

为$+\infty$, 而 $\int g^+ d\lambda = \int g^+ p d\mu$, $\int g^- d\lambda = \int g^- p d\mu$, 因而 $\int g d\lambda$ 存在的充分必要条件是 $\int g^+ p d\mu$ 与 $\int g^- p d\mu$ 不同时为 $+\infty$, 即 $\int g p d\mu$ 存在. 且在积分存在时

$$\begin{aligned}\int_A g d\lambda &= \int_A g^+ d\lambda - \int_A g^- d\lambda \\ &= \int_A g^+ p d\mu - \int_A g^- p d\mu \\ &= \int_A g p d\mu.\end{aligned}$$

最后对复可测函数 g, 只要分别考虑它的实部和虚部即可, 不再赘述. □

当我们把定理应用于

$$P_\xi(A) = \int_A \cdots \int p(x_1, \cdots, x_n) dx_1 \cdots dx_n.$$

P_ξ 相当于定理中的 λ, 而 μ 为 n 维 L 测度, 并应用 § 4 定理 4 的结论即得定理 6.

例 6　设 ξ 是在 $[a, b]$ 上均匀分布的随机变量, 由第二章知其分布密度为

$$p_\xi(x) = \begin{cases} 0, & x < a\text{或}x > b, \\ \dfrac{1}{b-a}, & a \leqslant x \leqslant b. \end{cases}$$

利用上面的公式, 得出

$$\begin{aligned}E\xi &= \int_{-\infty}^{\infty} x p(x) dx = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}, \\ D\xi &= E\left(\xi - \frac{a+b}{2}\right)^2 = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a} dx \\ &= \frac{1}{3}\left(\frac{b-a}{2}\right)^2 = \frac{1}{12}(b-a)^2.\end{aligned}$$

这里看到 ξ 的数学期望是 $[a, b]$ 的中点, 这同我们的直观理解是一致的.

从此例的简单计算中我们看到 R 积分在这里所起的作用. 当 R 积分存在时, 我们把积分看作 R 积分, 立刻可以应用数学分析的一切结果进行计算.

例 7　设 ξ 是按 $N(a, \sigma^2)$ 分布的正态随机变量, 它的分布密度为

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}},$$

我们来计算它的数学期望和方差.

$$E\xi=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}xe^{-\frac{(x-a)^2}{2\sigma^2}}dx$$
$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma t+a)e^{-\frac{t^2}{2}}dt,$$

其中 $t=\frac{x-a}{\sigma}$, 因为

$$\int_{-\infty}^{\infty}te^{-\frac{t^2}{2}}dt=0,\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}=\sqrt{2\pi},$$

故得

$$E\xi=a,$$
$$D\xi=E(\xi-a)^2=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-a)^2e^{-\frac{(x-a)^2}{2\sigma^2}}dx$$
$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sigma^2t^2e^{-\frac{t^2}{2}}dt\qquad\left(t=\frac{x-a}{\sigma}\right)$$
$$=\frac{\sigma^2}{\sqrt{2\pi}}\left\{\left[-te^{-\frac{t^2}{2}}\right]_{-\infty}^{+\infty}+\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt\right\}$$
$$=\sigma^2.$$

在上面的计算中由于被积函数是 R 绝对可积的, 因而可以看作 R 积分. 其中的变量替换既可看作 R 积分的变量替换又可看作 L 积分的变量替换.

由以上计算得知正态分布 $N(a,\sigma^2)$ 中两个参数分别为数学期望和方差. 因此正态分布律由数学期望和方差这两个数字特征完全决定. 因而 ξ 的高阶矩由 a,σ^2 表示. 通过计算求得 ξ 的 k 阶中心矩

$$\mu_k=E(\xi-a)^k=\begin{cases}0, & k\text{为奇数时},\\ \sigma^k\dfrac{k!}{2^{\frac{k}{2}}\left(\dfrac{k}{2}\right)!}, & k\text{为偶数时}.\end{cases}$$

对于按二维正态律 $N(a,b,\sigma_1^2,\sigma_2^2,r)$ 分布的随机变量 (ξ,η), 通过计算求得

$$E\xi=a,\quad E_\eta=b,$$
$$D\xi=\sigma_1^2,\quad D_\eta=\sigma_2^2,$$
$$b_{\xi\eta}=E(\xi-a)(\eta-b)=r\sigma_1\sigma_2,$$

$$r_{\xi\eta} = r.$$

这样, 二维正态律的五个参数 $a, b, \sigma_1^2, \sigma_2^2, r$ 分别是 ξ, η 的数学期望、方差和相关系数.

对于 n 维正态分布, 可求得类似的结果.

许多有用的分布律, 例如由按正态分布律分布的随机变量的函数产生的 χ^2-分布, χ-分布, t-分布, F-分布等等, 它们的分布密度及数字特征都可以应用这一节的公式计算出来, 请读者参看有关的概率论教科书. 这些分布在数理统计中起着重要的作用.

习题及补充

1. 设 ξ_1, ξ_2 是分别按具有参数 λ_1, λ_2 的 Poisson 律分布的独立随机变量. 求 $\eta = \xi_1 + \xi_2$ 的分布律.

2. 对 L 积分证明变量替换公式:若 $f(x)$ 是 $(R^{(1)}, \mathscr{B}^{(1)})$ 上的可测函数, $x = x(t)$ 在 $[\alpha, \beta]$ 上导数存在且大于零 (或小于零), 且 $x(t) = \int_\alpha^t x'(s)ds + x(\alpha)$, 若 $x(\beta) = b, x(\alpha) = a$, 则 $\int_a^b f(x)dx$ 存在的充分必要条件是 $\int_\alpha^\beta f(x(t))x'(t)dt$ 存在, 且

$$\int_a^b f(x)dx = \int_\alpha^\beta f(x(t))x'(t)dt.$$

(注意: 此处没有假定 $x'(s)$ 的连续性, 因而不能直接应用 § 5.2 的结果.)

3. 试对无界函数的 R-广义积分同 L 积分进行比较 (R- 广义积分定义见《微积分学教程》第三卷第一分册 590 目).

4. 设 ξ_1 与 ξ_2 是在 $[a, b]$ 上均匀分布的独立随机变量, 求 $\eta = \xi_1 + \xi_2$ 的分布函数及分布密度.

5. 设 ξ 是连续型随机变量, 分布函数为 $F(x)$, 试求 $F(\xi)$ 的分布律. 如果仅有 $F(x)$ 是连续函数, 结果如何?

6. 设 ξ, η 是独立随机变量, ξ 在 $[0, 1]$ 上均匀分布, η 按二项分布律 $B(n, p)$ 分布. 试证:$\xi + \eta$ 是连续型随机变量, 并求其分布密度.

7. 设 ξ, η 独立且都按 $N(0, \sigma^2)$ 分布, 令 $\zeta_1 = \alpha\xi + \beta\eta, \zeta_2 = \alpha\xi - \beta\eta$, 求 ζ_1 与 ζ_2 的相关系数 r_{ζ_1, ζ_2}.

8. 设 $\xi_1, \cdots, \xi_n$ 是具有相同分布密度 $p(x)$ 的独立随机变量, 且当 $x < 0$ 时 $p(x) = 0$, 当 $x \geqslant 0$ 时 $p(x)$ 连续, 其次设 ξ_k^* 为 $\xi_1, \cdots, \xi_n$ 按不降顺序排列时的第 k 个值, 试证 $(\xi_1^*, \cdots, \xi_r^*), 1 \leqslant r \leqslant n$ 的分布密度为

$$p(y_1,\cdots,y_r)=\begin{cases}\dfrac{n!}{(n-r)!}p(y_1)\cdots p(y_r)\left(\displaystyle\int_{y_r}^{\infty}p(x)dx\right)^{n-r}, 0\leqslant y_1\leqslant y_2\leqslant\cdots\leqslant y_r,\\ 0, \quad 其他.\end{cases}$$

(提示. 证明步骤如下:

i) 若有一 $y_k<0$, 则 $P(\xi_1^*<y_1,\cdots,\xi_r^*<y_r)=0$;

ii) 若 $y_k>y_{k+1}$, 则 $P(\xi_1^*<y_1,\cdots,\xi_r^*<y_r)=P(\cdots,\xi_k^*<y_{k+1},\xi_{k+1}^*<y_{k+1},\cdots)$;

iii) 若 $0\leqslant y_1\leqslant\cdots\leqslant y_r$, 则

$$\begin{aligned}&P(\xi_1^*<y_1,\cdots,\xi_r^*<y_r)\\ =&n!P(\xi_1<y_1,\cdots,\xi_r<y_r,\xi_1\leqslant\xi_2\leqslant\cdots\leqslant\xi_n)\\ =&\frac{n!}{(n-r)!}P(\xi_1<y_1,\cdots,\xi_r<y_r,\xi_1\leqslant\xi_2\leqslant\cdots\\ &\leqslant\xi_r,\xi_j\geqslant\xi_r,j=r+1,\cdots,n).\end{aligned}$$

将此式写成积分即可看出.)

9. 设 $F(x)$ 和 $G(x)$ 是两个有界分布函数, $f(x)$ 是连续函数, 且 $0<c_1\leqslant f(x)\leqslant c_2<+\infty(-\infty<x<\infty)$, 若对一切 $x\in R^{(1)}$,

$$F(x)=\int_{-\infty}^{x}f(y)dG(y),$$

则对一切 $y\in R^{(1)}$

$$G(y)=\int_{-\infty}^{y}\frac{1}{f(x)}dF(x).$$

(提示: 应用定理 7.)

10. 设 $\xi_1,\cdots,\xi_n$ 是按 $N(0,1)$ 分布的独立随机变量, 则随机变量

$$\chi^2=\sum_{k=1}^{n}\xi_k^2$$

的分布律称为**具有自由度为 n 的 χ^2- 分布**. 求 χ^2-分布的分布密度、数学期望及方差.

又设 $\chi=\sqrt{\chi^2}$, 试求 χ, $\dfrac{\chi}{\sqrt{n}}$ 的分布密度.

11. 设 ξ,η 是独立随机变量, 且 ξ 按 $N(0,1)$ 分布, $\eta=\dfrac{\chi}{\sqrt{n}}$, 则 $\zeta=\xi/\eta$ 的分布律称为自由度为 n 的 t-分布, 试求 t-分布的分布密度、数学期望及方差, $(n>1)$.

12. 设 $\chi^2_{n_1},\chi^2_{n_2}$ 是分别按自由度为 n_1,n_2 的 χ^2-分布的独立随机变量, 则称随机变量

$$F=\frac{\chi^2_{n_1}/n_1}{\chi^2_{n_2}/n_2}$$

为按第一、第二自由度分别为 n_1,n_2 的 F-分布的随机变量. 试求 F-分布的分布密度、数学期望和方差.

§3.6　r 次平均收敛与空间 L_r

r 次平均收敛是用 r 阶绝对矩 $E|\xi_n-\xi|^r$ 来度量偏差大小的一种收敛性. 当考察随机变量序列的 r 次平均收敛性时, 自然假定它们的 r 阶绝对矩有限. 因此这种研究是在 r 阶绝对矩有限的随机变量所成的类中进行的. 这个函数类按照一定方式定义距离, 作成一个距离空间. (关于距离空间的概念参看关肇直著《泛函分析讲义》第 1 章.) 即通常所谓的空间 L_r. 而 r 次平均收敛等价于按这个空间中的距离收敛. 所以在本节中我们也引出这种空间的定义和基本性质. 不过我们的注意力还是放在 r 次平均收敛的性质及 r 次平均收敛同几乎处处收敛、概率收敛、分布律收敛的相互关系上.

3.6.1　几个重要的不等式

1° 若 $a\geqslant 0,b\geqslant 0,0<\alpha<1,\alpha+\beta=1$, 则

$$a^\alpha b^\beta\leqslant \alpha a+\beta b, \tag{1}$$

且等号成立的充分必要条件是 $a=b$.

证　当 a,b 之一为零时, 结论显然成立. 故设 $ab\neq 0$, 只要证明

$$\left(\frac{a}{b}\right)^\alpha\leqslant \alpha\frac{a}{b}+\beta \tag{1$'$}$$

且等号成立当且仅当 $a=b$ 成立. 设 $x=\dfrac{a}{b}$, 则 $x>0$. 再设

$$f(x)=x^\alpha-\alpha x-\beta.$$

则

$$f'(x)=\alpha x^{\alpha-1}-\alpha\begin{cases}>0, & 0<x<1,\\ =0, & x=1,\\ <0, & x>1.\end{cases}$$

而当 $x=1$ 时 $f(x)=0$, 在 $0<x<1$ 上 $f(x)$ 递增, 在 $1<x<\infty$ 上 $f(x)$ 递减, 故

$$x^{\alpha}-\alpha x-\beta\leqslant 0,\quad 0<x<\infty,$$

当且仅当 $x=1$ 时等号成立. 这就证明了 $(1)'$, 并且当且仅当 $a=b$ 时 $(1)'$ 式中等号成立. □

2° **Hölder 不等式.** 设 $r>1, \dfrac{1}{r}+\dfrac{1}{s}=1$, 对任意随机变量 ξ,η 都有

$$E|\xi\eta|\leqslant E^{\frac{1}{r}}|\xi|^r E^{\frac{1}{s}}|\eta|^s, \tag{2}$$

其中 $E^{\frac{1}{r}}|\xi|^r$ 表示 $(E|\xi|^r)^{\frac{1}{r}}$ (以后经常应用类似记号, 不再一一声明). 若 (2) 中右端有限, 则等号成立的充分必要条件是存在不全为零的常数 c_1,c_2 使 $c_1|\xi|^r+c_2|\eta|^s=0$, a.e. 成立.

证 当 $\xi=0$, a.e. 或 $\eta=0$, a.e. 或 $E^{\frac{1}{r}}|\xi|^r=\infty$ 或 $E^{\frac{1}{s}}|\eta|^s=\infty$ 时 (2) 式显然成立, 今设 $0<E|\xi|^r, E|\eta|^s<\infty$, 令

$$a=\frac{|\xi|^r}{E|\xi|^r}, b=\frac{|\eta|^s}{E|\eta|^s}, \alpha=\frac{1}{r}, \beta=\frac{1}{s}.$$

利用不等式 (1) 即得

$$\frac{|\xi\eta|}{E^{\frac{1}{r}}|\xi|^r E^{\frac{1}{s}}|\eta|^s}\leqslant\frac{1}{r}\frac{|\xi|^r}{E|\xi|^r}+\frac{1}{s}\frac{|\eta|^s}{E|\eta|^s}. \tag{2$'$}$$

两边取数学期望, 即得 Hölder 不等式. 显然, (2) 中等号成立的充分条件是 $(2)'$ 中等号 a.e. 成立. 反之, 如果 $(2)'$ 在某一非零概率集 A 上等号不成立, 则 (2) 中等号不能成立. 由 1° 易得等号成立的充要条件. □

当 $r=2$ 时 Hölder 不等式即是熟知的 Schwarz 不等式,

$$|E\xi\bar{\eta}|^2\leqslant E|\xi|^2\cdot E|\eta|^2.$$

3° 关于数列的 C_r-不等式.

$$|a_1+\cdots+a_n|^r\leqslant C_r|a_1|^r+\cdots+C_r|a_n|^r, \tag{3}$$

其中

$$C_r=\begin{cases}1, & 0<r\leqslant 1,\\ n^{r-1}, & r>1.\end{cases}$$

(3) 中等号成立的充要条件, 当 $r>1$ 时是 $a_1=\cdots=a_n$; 当 $r=1$ 时是 $a_1,\cdots,a_n$ 同号, 当 $r<1$ 时是 $a_1,\cdots,a_n$ 中至多有一个不为零.

证　$r>1$ 的情形: 设 ξ 是以概率 $\frac{1}{n}$ 分别取 $a_1,\cdots,a_n$ 为值的随机变量, 则

$$E|\xi|=\frac{1}{n}(|a_1|+\cdots+|a_n|),$$

$$E|\xi|^r=\frac{1}{n}(|a_1|^r+\cdots+|a_n|^r).$$

利用 Hölder 不等式 (取 $\eta=1$),

$$\frac{1}{n^r}(|a_1|+\cdots+|a_n|)^r\leqslant\frac{1}{n}(|a_1|^r+\cdots+|a_n|^r).$$

由此即得 (3) 式. 在上式中等号成立的充分必要条件为 $|a_1|=\cdots=|a_n|$, 再由绝对值的性质知 (3) 中等号成立的充分必要条件为 $a_1=\cdots=a_n$.

$r\leqslant 1$ 的情形: 只需证明 $a_1,\cdots,a_n$ 不全为零的情形. 由于 $r\leqslant 1$

$$\frac{|a_k|}{|a_1|+\cdots+|a_n|}\leqslant\frac{|a_k|^r}{(|a_1|+\cdots+|a_n|)^r},\quad k=1,\cdots,n.$$

将上述不等式相加即得 (3) 式.

当 $r=1$ 时, 由绝对值的性质知 (3) 中等号成立的充分必要条件为 $a_1,\cdots,a_n$ 同号.

当 $r<1$ 时等号成立的充分必要条件是 $a_1,\cdots,a_n$ 中至多有一个不为零.

4° C_r-不等式. 设 $\xi_1,\cdots,\xi_n$ 是随机变量, 则

$$E|\xi_1+\cdots+\xi_n|^r\leqslant C_rE|\xi_1|^r+\cdots+C_rE|\xi_n|^r,\tag{4}$$

其中 C_r 如 3° 中所述. 由 3° 易得等号成立的条件.

由 C_r-不等式容易看出, 若 $\xi_1,\cdots,\xi_n$ 的 r 阶绝对矩有限, 则它们的和 $\xi_1+\cdots+\xi_n$ 的 r 阶绝对矩也有限.

5° **Minkowski 不等式.**　设 $r\geqslant 1,\xi,\eta$ 是随机变量, 则

$$E^{\frac{1}{r}}|\xi+\eta|^r\leqslant E^{\frac{1}{r}}|\xi|^r+E^{\frac{1}{r}}|\eta|^r.\tag{5}$$

等号成立的充分必要条件当 $r>1$ 时是存在不全为零且同号的 c_1,c_2 使 $c_1\xi-c_2\eta=0$, a.e. 成立, 当 $r=1$ 时是 ξ,η, a.e. 同号.

证　$r=1$ 时显然. $r>1$ 时利用 Hölder 不等式,

$$\begin{aligned}E|\xi+\eta|^r&\leqslant E|\xi||\xi+\eta|^{r-1}+E|\eta||\xi+\eta|^{r-1}\\&\leqslant E^{\frac{1}{r}}|\xi|^r.E^{1-\frac{1}{r}}|\xi+\eta|^{\frac{r-1}{1-\frac{1}{r}}}+E^{\frac{1}{r}}|\eta|^r\cdot E^{1-\frac{1}{r}}|\xi+\eta|^{\frac{r-1}{1-\frac{1}{r}}}\\&=(E^{\frac{1}{r}}|\xi|^r+E^{\frac{1}{r}}|\eta|^r)E^{1-\frac{1}{r}}|\xi+\eta|^r.\end{aligned}$$

上式等号成立的充分必要条件是 ξ,η, a.e. 同号且存在不全为零的 c_1, c_2 及 c_1', c_2' 使 $c_1|\xi|^r + c_2|\xi+\eta|^r = 0$, a.e. 且 $c_1'|\eta|^r + c_2'|\xi+\eta|^r = 0$, a.e. 由此易得关于等号成立的结论.

如果 $E|\xi+\eta|^r = 0$, (5) 显然成立, 如果 $E|\xi+\eta|^r = \infty$, 由 (4) 知 $E|\xi|^r = \infty$ 或 $E|\eta|^r = \infty$, 因而 (5) 也成立. 而 $0 < E|\xi+\eta|^r < \infty$ 时, 由上式即得 (5) 式.

6° 如果 $E|\xi|^r$ 有限, 则

$$E^{\frac{1}{r'}}|\xi|^{r'} \leqslant E^{\frac{1}{r}}|\xi|^r, \quad 当 0 < r' < r. \tag{6}$$

证 在 Hölder 不等式中取 r 为 $\dfrac{r}{r'} > 1$, 取 ξ 为 $|\xi|^{r'}$, 取 $\eta = 1$ 即得所欲证.

3.6.2 L_r 空间

定理 1 设 L_r 是概率场 $(\Omega, \mathscr{A}, P)$ 上一切 r 阶绝对矩有限的随机变量所成的类, 则 L_r 对于如下定义的距离

$$d(\xi,\eta) = \begin{cases} E|\xi-\eta|^r, & 0 < r \leqslant 1, \quad \xi,\eta \in L_r, \\ E^{\frac{1}{r}}|\xi-\eta|r, & r > 1, \end{cases}$$

作成一个距离空间 (a.e. 相等的随机变量看成 L_r 中同一元素).

证. 当 $\xi,\eta \in L_r$ 时, $E|\xi|^r, E|\eta|^r$ 是有限实数, 根据 G_r 不等式 $E|\xi-\eta|^r$ 也是有限实数, 可见上面定义的 $d(\xi,\eta)$ 是一有限实数. 现在验证 $d(\xi,\eta)$ 是 L_r 中的距离:

i) $d(\xi,\eta) \geqslant 0$, 对任意 $\xi,\eta \in L_r$;

ii) $d(\xi,\eta) = 0$ 当且仅当 $\xi = \eta$, a.e.;

iii) $d(\xi,\eta) = d(\eta,\xi)$, 对任意 $\xi,\eta \in L_r$ 成立.

以上三点显然成立. 至于

iv) 三角不等式: 对于任意 $\xi,\eta,\zeta \in L_r$,

$$d(\xi,\eta) \leqslant d(\xi,\zeta) + d(\zeta,\eta),$$

当 $0 < r \leqslant 1$ 时就是 C_r-不等式, 当 $r > 1$ 时就是 Minkowski 不等式. □

定义 1 如果 $\xi_n, \xi \in L_r, n = 1,2,\cdots, E|\xi_n - \xi|^r \to 0 (n \to \infty)$, 则称 $\{\xi_n\}$ ***r* 次平均收敛** 于 ξ, 记作 $\xi_n \xrightarrow{r} \xi$. 当 $r = 2$ 时, 通常也称为 $\{\xi_n\}$**均方收敛**于 ξ, 记作 $\xi_n \xrightarrow{\text{q.m.}} \xi$ 或 $\underset{n\to\infty}{\text{l.i.m}}\, \xi_n = \xi$.

显然, r 次平均收敛等价于按 L_r 的距离收敛.

其次, 当 $\xi_n \in L_r, E|\xi_n - \xi|^r \to 0$ 时, 利用 C_r-不等式

$E|\xi|^r \leqslant C_r E|\xi_n|^r + C_r E|\xi_n - \xi|^r < \infty$, ($n$ 充分大). 因而 $\xi \in L_r, \xi_n \xrightarrow{r} \xi$. 所以

为了证明 $\xi_n \xrightarrow{r} \xi$, 当已知 $\xi_n \in L_r, n = 1, 2, \cdots$ 时, 只需验证 $E|\xi_n - \xi|^r \to 0$, 而不必预先验证 $\xi \in L_r$.

定理 2　设 $\{\xi_n\} \subset L_r, \xi_n \xrightarrow{r}$ 某一 ξ 的充分必要条件是 $E|\xi_m - \xi_n|^r \to 0 (m, n \to \infty)$, 因而 L_r 是一完备距离空间.

证　条件的必要性. 假定 $\xi_n \xrightarrow{r} \xi$, 利用 C_r-不等式,

$$E|\xi_m - \xi_n|^r \leqslant C_r E|\xi_m - \xi|^r + C_r E|\xi_n - \xi|^r \to 0, (m, n \to \infty).$$

条件的充分性. 假定 $E|\xi_m - \xi_n|^r \to 0 (m, n \to \infty)$ 根据Чебышев 不等式 (§ 2.2 定理 2 推论), 对于任意 $\varepsilon > 0$,

$$P\{|\xi_m - \xi_n| \geqslant \varepsilon\} \leqslant \frac{1}{\varepsilon^r} E|\xi_m - \xi_n|^r \to 0, (m, n \to \infty).$$

即 $\{\xi_n\}$ 依概率相互收敛, 因而根据第二章 §5 定理 5 存在子序列 $\xi_{n'} \xrightarrow{\text{a.e.}}$ 某一 ξ. 今证 $E|\xi_m - \xi|^r \to 0$.

对于任一固定的 m,

$$\xi_m - \xi_{n'} \xrightarrow{\text{a.e.}} \xi_m - \xi, (n' \to \infty).$$

利用 Fatou-Lebesgue 定理 (§ 3 定理 1)

$$E|\xi_m - \xi|^r = E \varliminf_{n' \to \infty} |\xi_m - \xi_{n'}|^r \leqslant \varliminf_{n' \to \infty} E|\xi_m - \xi_{n'}|^r.$$

令 $m \to \infty$, 取上极限,

$$\varlimsup_{m \to \infty} E|\xi_m - \xi|^r \leqslant \varlimsup_{m \to \infty} \varliminf_{n' \to \infty} E|\xi_m - \xi_{n'}|^r.$$

由于 $\lim\limits_{\substack{m \to \infty \\ n' \to \infty}} E|\xi_m - \xi_{n'}|^r = 0$ 等价于

$$\lim_{m \to \infty} E|\xi_m - \xi_{m+v'}|^r = 0, \text{对} v' \text{一致地成立},$$

因而 $\lim\limits_{m \to \infty} \varliminf\limits_{v' \to \infty} E|\xi_m - \xi_{m+v'}|^r = 0$ 即

$$\lim_{m \to \infty} \varliminf_{n' \to \infty} E|\xi_m - \xi_{n'}|^r = 0.$$

故

$$\varlimsup_{m \to \infty} E|\xi_m - \xi|^r = 0.$$

由于非负序列下极限非负, 故 $\lim\limits_{m \to \infty} E|\xi_m - \xi|^r = 0$.　□

L_r 总是由一给定的概率的空间 $(\Omega,\mathscr{A},P)$ 上的具有有限的 r 阶绝对矩的随机变量组成的,所以确切地应记作 $L_r(\Omega,\mathscr{A},P)$. 常常也要考虑一般测度空间 $(\Omega,\mathscr{A},\mu)$ 上的 r 次绝对可积函数 $\left(\text{即} \int|f|^r d\mu<\infty \text{ 的函数 } f\right)$ 所成的空间 $L_r(\Omega,\mathscr{A},\mu)$ (不过通常多限于 $r\geqslant 1$ 的情形). L_r 中的距离永远指的是按定理 1 中的方式所定义的. 应该指出当 $r\geqslant 1$ 时 L_r 中的距离可由 L_r 中的范数 $\|\xi\|=E^{\frac{1}{r}}|\xi|^r$ 导出, 由于 L_r 对于这个距离是完备的, 所以 $L_r(r\geqslant 1)$ 是 Banach 空间, 当 $r<1$ 时, L_r 中的距离则由 L_r 中的准范数 $\|\xi\|=E|\xi|^r$ 导出, 因而 $L_r(0<r<1)$ 只是一个 Frechet 空间 (参看文献 [25] 第 2 章).

由 (6) 式知当 $E|\xi|^r<\infty$ 时, $E|\xi|^{r'}<\infty, 0<r'<r$, 所以函数类 L_r 与 $L_{r'}$ 之间有下列的包含关系:

$$L_{r'}\supset L_r,\quad (0<r'<r).$$

由于 $L_r, l_{r'}$ 中的距离的定义不同, 上面的包含关系并不意味着 L_r 是 $L_{r'}$ 的子距离空间. 但 r' 次与 r 次平均收敛之间具有下列关系:

性质 1 若 $\xi_n\xrightarrow{r}\xi, 0<r'<r$, 则 $\xi_n\xrightarrow{r'}\xi$.

关于 r 次平均收敛, 我们还有

性质 2 若 $\xi_n\xrightarrow{r}\xi$, 则 $E|\xi_n|^r\to E|\xi|^r$.

证 $r\leqslant 1$ 的情形: 利用 C_r 不等式,

$$E|\xi_n|^r\leqslant E|\xi|^r+E|\xi_n-\xi|^r,$$

$$E|\xi|^r\leqslant E|\xi_n|^r+E|\xi_n-\xi|^r,$$

故

$$|E|\xi_n|^r-E|\xi|^r|\leqslant E|\xi_n-\xi|^r\to 0.$$

$r\geqslant 1$ 的情形: 仿照上面的方法, 利用 Minkowski 不等式,

$$|E^{\frac{1}{r}}|\xi_n|^r-E^{\frac{1}{r}}|\xi|^r|\leqslant E^{\frac{1}{r}}|\xi_n-\xi|^r\to 0. \qquad \square$$

性质 2 也可由距离空间的性质直接得到.

3.6.3 r 次平均收敛与各种收敛性之间的关系

现在我们来比较各种收敛性之间的关系, 为此我们引出下面的

定义 2 设 $\xi_t, t\in T$ 是一族随机变量, 其中 T 是任一指标集.

i) 如果当 $P(A)\to 0$ 时 $\int_A|\xi_t|dP\to 0$ 对 $t\in T$ 一致成立, 即对任一 $\varepsilon>0$, 存在 $\delta_\varepsilon>0$, 使当 $P(A)<\delta_\varepsilon$ 时, $\int_A|\xi_t|dP<\varepsilon$, 对一切 $t\in T$ 成立, 则称 $\xi_t, t\in T$的积

分一致连续或称 $\int_A |\xi_t|dP, t\in T$ **一致连续**.

ii) 如果当 $a\to\infty$ 时,

$$\int_{\{|\xi_t|\geqslant a\}} |\xi_t|dP \to 0, \text{对} t\in T \text{一致成立},$$

则称 $\xi_t, t\in T$ **一致可积**.

iii) 如果存在一个常数 c , 使得

$$\int |\xi_t|dP \leqslant c, t\in T,$$

则称 $\xi_t, t\in T$**的积分一致有界**或 $\int |\xi_t|dP$ **一致有界**.

引理 1　随机变量族 $\xi_t, t\in T$ 一致可积的充分必要条件是 $\xi_t, t\in T$ 的积分一致有界且一致连续.

证　必要性　假定 $\xi_t, t\in T$ 一致可积, 则对于任一 $\varepsilon>0$, 存在一个常数 c, 使得

$$\int_{\{|\xi_t|\geqslant c\}} |\xi_t|dP < \varepsilon/2, \text{对一切} t\in T.$$

于是当 $P(A)<\dfrac{\varepsilon}{2c}$ 时,

$$\begin{aligned}\int_A |\xi_t|dP &= \int_{A\cap\{|\xi_t|<c\}} |\xi_t|dP + \int_{A\cap\{|\xi_t|\geqslant c\}} |\xi_t|dP \\ &\leqslant cP(A)+\frac{\varepsilon}{2} < \varepsilon, \text{对一切} t\in T.\end{aligned}$$

即 $\xi_t, t\in T$ 的积分一致连续. 又

$$\int |\xi_t|dP = \int_{\{|\xi_t|<c\}} |\xi_t|dP + \int_{\{|\xi_t|\geqslant c\}} |\xi_t|dP < c+\frac{\varepsilon}{2},$$

对一切 $t\in T$. 故知 $\xi_t, t\in T$ 的积分一致有界.

充分性　假定 $\xi_t, t\in T$ 的积分一致有界且一致连续. 由一致连续性知, 对于任意 $\varepsilon>0$ 存在一 $\delta_\varepsilon>0$ (与 t 无关), 使得当 $P(A)<\delta_\varepsilon$ 时,

$$\int_A |\xi_t|dP < \varepsilon, \text{对一切} t\in T;$$

利用积分的一致有界性, 设 $\int |\xi_t|dP < c, t\in T$, 以及 § 2 定理 2 推论, 得

$$P(|\xi_t|\geqslant a) \leqslant \frac{1}{a}\int |\xi_t|dP < \frac{c}{a}.$$

所以当 $a > a_0 = c/\delta_\varepsilon$(与 t 无关) 时,

$$P(|\xi_t| \geqslant a) < \delta_\varepsilon,$$

因而当 $a > a_0$ 时,

$$\int_{\{|\xi_t|\geqslant a\}} |\xi_t| dP < \varepsilon, \text{对一切} t \in T,$$

即 $\xi_t, t \in T$ 一致可积.

引理 2 若 $\xi_n \xrightarrow{r} \xi$, 则 $|\xi_n|^r$ 一致可积.

证 对 $\xi_n\chi_A, \xi\chi_A$ 利用 C_r-不等式,

$$\begin{aligned}\int_A |\xi_n|^r &\leqslant C_r \int_A |\xi|^r + C_r \int_A |\xi_n - \xi|^r \\ &\leqslant C_r \int_A |\xi|^r + C_r \int |\xi_n - \xi|^r.\end{aligned}$$

由于 $\xi_n \xrightarrow{r} \xi$, 当 $n > n_0$ (与 A 无关) 时 $\int |\xi_n - \xi|^r < \frac{\varepsilon}{2C_r}$, 而由 $\int_A |\xi|^r dP \longrightarrow 0$ (当 $P(A) \to 0$ 时) 知存在 δ_1 , 使当 $P(A) < \delta_1$ 时 $\int_A |\xi|^r dP < \frac{\varepsilon}{2C_r}$, 由此可见, 当 $P(A) < \delta_1$ 时

$$\int_A |\xi_n|^r dP < \varepsilon, n > n_0.$$

对于前面有限个积分 $\int_A |\xi_n|^r dP, n = 1, \cdots, n_0$, 由于当 $P(A) \to 0$ 时, 它们分别以零为极限 (见 § 3 推论 3) 所以存在 δ_2, 使当 $P(A) < \delta_2$ 时

$$\int_A |\xi_n|^r dP < \varepsilon, n = 1, \cdots, n_0,$$

于是当 $P(A) < \delta = \min(\delta_1, \delta_2)$ 时, $\int_A |\xi_n|^r dP < \varepsilon$, 对一切 $n \geqslant 1$ 成立. 即 $\int_A |\xi_n|^r dP, n = 1, 2, \cdots$ 一致连续. 同时由

$$\int |\xi_n|^r dP \leqslant c_r \int |\xi|^r dP + c_r \int |\xi_n - \xi|^r dP.$$

不难看出 $\int |\xi_n|^r dP$ 一致有界, 因而利用引理 1, $|\xi_n|^r, n = 1, 2, \cdots$ 一致可积. □

应用相同的证明方法可以将引理 2 推广成下列的

推论 若 $\xi_t \xrightarrow{r} \xi$, 当 $t \to t_0$ (即 $\lim\limits_{t\to t_0} E|\xi_t - \xi|^t = 0$), 则当 $|t_0| < \infty(t_0 = \infty$或$t_0 = -\infty)$ 时有一正数 t_1 存在使得 $|\xi_t|^r, |t - t_0| \leqslant t_1(t \geqslant t_1$或$t < -t_1)$ 一致可

积.

定理 3　$\{\xi_n\} \subset L_r$, 则 $\xi_n \xrightarrow{r} \xi$ 的充分必要条件是

i) $\xi_n \xrightarrow{P} \xi$ 且 $\int_A |\xi_n - \xi|^r dP$ 一致连续,

或

ii) $\xi_n \xrightarrow{P} \xi$ 且 $\int_A |\xi_n|^r$ 一致连续.

证　首先假设 $\xi_n \xrightarrow{r} \xi$, 利用 § 2 定理 2 推论

$$P\{|\xi_n - \xi| \geqslant \varepsilon\} \leqslant \frac{1}{\varepsilon^r} E|\xi_n - \xi|^r \to 0,$$

即 $\xi_n \xrightarrow{P} \xi$.

其次, 由 $\xi_n \xrightarrow{r} \xi$ 知 $\xi_n - \xi \xrightarrow{r} 0$, 由引理 2 和引理 1 知 $\int_A |\xi_n - \xi|^r dP$ 及 $\int_A |\xi_n|^r dP, n = 1, 2, \cdots$ 一致连续. 故条件 i) 及 ii) 的必要性获证.

现在证明条件 i) 的充分性. 已知 $\xi_n \xrightarrow{P} \xi, \int_A |\xi_n - \xi|^r dP$ 一致连续, 故任给 $\varepsilon > 0$, 存在 $\delta_\varepsilon > 0$, 使当 $P(A) < \delta_\varepsilon$ 时, $\int_A |\xi_n - \xi|^r dP < \varepsilon$, 而由 $\xi_n \xrightarrow{P} \xi$, 存在 n_0, 当 $n > n_0$ 时, $P\{|\xi_n - \xi| \geqslant \varepsilon\} < \delta_\varepsilon$, 于是当 $n > n_0$ 时,

$$E|\xi_n - \xi|^r = \int_{\{|\xi_n - \xi| \geqslant \varepsilon\}} |\xi_n - \xi|^r dP + \int_{\{|\xi_n - \xi| < \varepsilon\}} |\xi_n - \xi|^r dP < \varepsilon + \varepsilon^r,$$

即 $E|\xi_n - \xi|^r \to 0, \xi_n \xrightarrow{r} \xi$.

最后证明条件 ii) 的充分性. 为此只需证 ii) 蕴含 i). 由于 $\xi_n \xrightarrow{P} \xi$, 存在一子序列 $\xi_{n'} \xrightarrow{\text{a.e.}} \xi$, 因而 $|\xi_{n'}|^r \chi_A \xrightarrow{\text{a.e.}} |\xi|^r \chi_A$, 利用 Fatou-Lebesgue 定理,

$$\begin{aligned}\int_A |\xi|^r dP = \int |\xi|^r \chi_A dP &\leqslant \varliminf_{n' \to \infty} \int |\xi_{n'}|^r \chi_A dP \\ &= \varliminf_{n' \to \infty} \int_A |\xi_{n'}|^r dP.\end{aligned}$$

由于 $\int_A |\xi_n|^r$ 一致连续, 任给 $\varepsilon > 0$, 存在 $\delta > 0$, 使当 $P(A) < \delta$ 时, $\int_A |\xi_n|^r dP < \varepsilon$, 故 $\varliminf_{n' \to \infty} \int_A |\xi_{n'}|^r < \varepsilon$, 因而当 $P(A) < \delta$ 时

$$\int_A |\xi|^r < \varepsilon,$$

这样一来, 利用 C_r-不等式及 $\int_A |\xi_n|^r dP \ n=1,2,\cdots$ 的一致连续性, 当 $P(A)<\delta$ 时,

$$\int_A |\xi_n-\xi|^r dP \leqslant C_r\int_A |\xi_n|^r dP + C_r\int_A |\xi|^r dP < 2C_r\varepsilon,$$

即 $\int_A |\xi_n-\xi|^r dP$ 一致连续. 于是 ii) 的充分性获证. □

推论 1 若 $\xi_n \xrightarrow{r} \xi$, 则 $\xi_n \xrightarrow{P} \xi$.

推论 2 若 $\sup\limits_n E|\xi_n|^r = c < \infty, \xi_n \xrightarrow{P} \xi$, 则当 $0<r'<r$ 时, $\xi_n \xrightarrow{r'} \xi$.

证 由于对任意 $a>0$,

$$\begin{aligned}\int_A |\xi_n|^{r'} &= \int_{A\cap\{|\xi_n|\geqslant a\}} |\xi_n|^{r'} + \int_{A\cap\{|\xi_n|<a\}} |\xi_n|^{r'} \\ &\leqslant \int_{A\cap\{|\xi_n|\geqslant a\}} |\xi_n|^{r'}\frac{|\xi_n|^{r-r'}}{a^{r-r'}} + a^{r'}P(A) \\ &\leqslant ca^{r'-r} + a^{r'}P(A).\end{aligned}$$

任给 $\varepsilon>0$, 取充分大的 a , 使 $ca^{r'-r}<\frac{\varepsilon}{2}$, 于是当 $P(A)<\frac{\varepsilon}{2a^{r'}}$ 时, 可使

$$\int_A |\xi_n|^{r'} < \varepsilon, 对一切n,$$

即 $\int_A |\xi_n|^{r'}, n=1,2,\cdots$ 一致连续. 因此根据定理 3 知 $\xi_n \xrightarrow{r'} \xi$. □

推论 3 如果 $|\xi_n|\leqslant \eta\in L_r, \xi_n \xrightarrow{p} \xi$, 则 $\xi_n \xrightarrow{r} \xi$.

证 $|\xi_n|^r \leqslant |\eta|^r$, 故 $\xi_n\in L_r$, 且 $\int_A |\xi_n|^r \leqslant \int_A |\eta|^r$, 因而 $\int_A |\xi_n|^r, n=1,2,\cdots$ 一致连续. 由定理 3, $\xi_n \xrightarrow{r} \xi$. □

推论 4 若 $\{\xi_n\}$ 一致有界 (即 $|\xi_n|\leqslant c$(常数)$, n=1,2,\cdots), \xi_n \xrightarrow{p} \xi$, 则 $\xi_n \xrightarrow{r} \xi$.

这是推论 3 当 $\eta=c$ 的情形.

应用与证明定理 3 相同的方法, 还可以将定理 3 推广成下述

推论 5 设 $\xi_t\in L_r, t\in T\subset R^{(1)}$, 则 $\xi_t \xrightarrow{r} \xi, (t\to t_0)$ 的充分与必要条件是

i) $\xi_t \xrightarrow{P} \xi(t\to t_0)$ 且有一 $t_1>0$ 存在, 使当 t_0 有限时, $\int_A |\xi_t-\xi|^r, |t-t_0|\leqslant t_t$ 一致连续; 而当 $t_0=\infty$ 时, $\int |\xi_t-\xi|^r, t\geqslant t_1$ 一致连续; 当 $t_0=-\infty$ 时, $\int_A |\xi_t-\xi|^r, t\leqslant -t_1$ 一致连续,

或

ii) $\xi_t \xrightarrow{P} \xi(t \to t_0)$ 且有一 $t_1 > 0$ 存在, 使当 t_0 有限时, $\int_A |\xi_t|^r, |t - t_0| \leqslant t_1$ 一致连续; 当 $t_0 = \infty$ 时, $\int_A |\xi_t|^r, t \geqslant t_1$ 一致连续; 当 $t_0 = -\infty$ 时 $\int_A |\xi_t|^r, t \leqslant -t_1$ 一致连续.

总结起来, 各种收敛性的关系可列成下表

$$\begin{array}{ccccc} \xi_n \xrightarrow{\text{a. e.}} \xi & \Longrightarrow & \xi_n \xrightarrow{P} \xi & \Longrightarrow & \mathscr{L}(\xi_n) \to \mathscr{L}(\xi) \\ & \swarrow\!\!\!\swarrow & \Uparrow & & \\ \xi_{n_k} \xrightarrow{\text{a. e.}} \xi & & \xi_n \xrightarrow{r} \xi & \Longrightarrow & \xi_n \xrightarrow{r'} \xi \ (0 < r' < r) \end{array}$$

习题及补充

1. 设 ξ 是任意的实随机变量, $g(x)$ 是 $R^{(1)}$ 上非负 Borel 可测函数.

i) 若 g 是偶函数且在 $[0, +\infty)$ 上非降, 则对于每个 $a \geqslant 0$,

$$\frac{Eg(\xi) - g(a)}{\text{a.e. sup } g(\xi)} \leqslant P(|\xi| \geqslant a) \leqslant \frac{Eg(\xi)}{g(a)},$$

其中 a.e. sup $g(\xi) = \inf\{c: \ P(g(\xi) \geqslant c) = 0\}$.

ii) 若 g 在 R 上非降, 则上式中间的一项应换为 $P(\xi \geqslant a)$, 此处 a 为任一实数.

iii) 在 i) 中令 $g(x) = |x|^r (r > 0)$, 将 ξ 换作 $\xi_n - \xi$, 证明: 若 $\xi_n \xrightarrow{r} \xi$, 则 $\xi_n \xrightarrow{P} \xi$;反之, 若 ξ_n a.e. 一致有界且 $\xi_n \xrightarrow{P} \xi$, 则 $\xi_n \xrightarrow{r} \xi$.

iv) 在 i) 中令 $g(x) = \dfrac{|x|^r}{1 + |x|^r} (r > 0)$, 则

$$\xi_n \xrightarrow{P} \xi \Leftrightarrow E\frac{|\xi_n - \xi|^r}{1 + |\xi_n - \xi|^r} \to 0.$$

取 $r = 1$, 令 $d(\xi, \eta) = E\dfrac{|\xi - \eta|}{1 + |\xi - \eta|}$, 把 a.e. 相等的随机变量看成同一元素, 证明在$(\Omega, \mathscr{A}, P)$ 上随机变量按距离 $d(\xi, \eta)$ 所作成的空间是完备距离空间. (证明 $d(\xi, \eta)$ 满足定理 1 中的四条性质, 且基本列收敛.)

2. 利用概率方法证明不等式: 对 $r \geqslant 1, x_i, y_i$ 有限, $i = 1, \cdots, n$,

$$\left(\sum_{i=1}^n |x_i + y_i|^r\right)^{\frac{1}{r}} \leqslant \left(\sum_{i=1}^n |x_i|^r\right)^{\frac{1}{r}} + \left(\sum_{i=1}^n |y_i|^r\right)^{\frac{1}{r}}.$$

3. 证明Ляпунов 不等式，设 $\mu_r = E|\xi|^r$, 如果 $r \geqslant s \geqslant t \geqslant 0$, 则有 $\mu_r^{s-t}\mu_s^{t-r}\mu_t^{r-s} \geqslant 1$.

4. 若 $(\Omega, \mathscr{A}, \mu)$ 是测度空间, 可以相应地构造一个距离空间 $(\mathscr{X}, d)$: 将 $\mathscr{A}$ 中所有具有有限测度的集合 $A, B, \cdots$ 构成的空间取作 $\mathscr{X}$, 将 $\mu(A\Delta B) = 0$ 的集 A, B 视为同一元素, 按 $d(A, B) = \mu(A\Delta B)$ 规定距离, 试证此距离空间是完备的.

(提示: 若 $\mu(A_n\Delta A_m) \to 0(n, m \to \infty)$, 则 $\chi_{A_n} - \chi_{A_m} \xrightarrow{\mu} 0(n, m \to \infty)$, 因而 $\{\chi_{A_n}\}$ 依测度收敛.)

5. 称定义在有限或无限区间 $I \subset R$ 上的实函数 g 为凸函数, 如果对一切 $x, x' \in I$ 有

$$g\left(\frac{x + x'}{2}\right) \leqslant \frac{1}{2}g(x) + \frac{1}{2}g(x').$$

对于 N 维区间 I^N 上的函数 g, 亦可给出同样定义. 并且在 N 维情形下 $g(x)$ 是凸函数等价于作为单变数 u 的函数 $f(u) = g(x + ux')$ 是凸函数. 试证:

i) g 是 I 上连续凸函数当且仅当对一切 $x_0 \in I$, 存在 $\lambda(x_0)$, 使当 $x \in I$ 时恒有

$$\lambda(x_0)(x - x_0) \leqslant g(x) - g(x_0);$$

ii) 设 ξ 是一随机变量, $E\xi$ 有限, 且 $\xi(\omega) \in I$, 对几乎所有的 ω 成立. g 是 I 上的凸函数, 则有

$$g(E\xi) \leqslant Eg(\xi).$$

6. 若 $\xi_n \xrightarrow{r} \xi, \eta_n \xrightarrow{s} \eta, \dfrac{1}{r} + \dfrac{1}{s} = 1, r > 1$, 则 $\xi_n\eta_n \xrightarrow{1} \xi\eta$.

7. 设 $\{f_n\}$ 是 $(\Omega, \mathscr{A}, \mu)$ 上的可测函数列, 且 $|f_n| \leqslant g, g$ 关于 μ 可积, 若 $f_n \xrightarrow{\text{a.e.}} f$ 或 $f_n \xrightarrow{\mu} f$ 则 $f_n \xrightarrow{1} f$.

8. i) $\sum|\xi_n| < \infty$, a.e., 当且仅当部分和序列的分布函数收敛向一个概率分布函数 (在它的连续点上).

ii) 若 $\sum E|\xi_n|^r < \infty(r < 1)$ 或 $\sum E^{\frac{1}{r}}|\xi_n|^r < \infty(r \geqslant 1)$, 则 $\sum|\xi_n| < \infty$, a.e..

§3.7 不定积分与 σ-可加集函数的分解

在 §3 中我们看到不定积分是给定空间 $(\Omega, \mathscr{A})$ 上的 σ-可加集函数. 本节将要讨论任一 σ-可加集函数可以表成关于这个空间上给定测度 μ 的不定积分的条件. 所得到的 Radon-Nikodym 定理在概率论中有重要意义. 另外我们将证明每一分布函数都可以表成三个分布函数, 即绝对连续部分、离散部分和奇异部分的和.

所有这些讨论都基于 σ-可加集函数的分解定理, 所以我们先来讨论 σ-可加集函数的性质.

3.7.1　σ-可加集函数的分解定理

定理 1　设 φ 是 $(\Omega, \mathscr{A})$ 上的一个实值 σ-可加集函数, 则存在 $C, D \in \mathscr{A}$, 使

$$\varphi(C) = \sup_{A \in \mathscr{A}} \varphi(A), \quad \varphi(D) = \inf_{A \in \mathscr{A}} \varphi(A), \tag{1}$$

即 φ 在 $\mathscr{A}$ 上达到极大值与极小值.

证　由于证明的类似, 只证 C 的存在性.

如果存在 $A \in \mathscr{A}$, 使 $\varphi(A) = +\infty$, 则取 $C = A$ 即可.

假设对任何 $A \in \mathscr{A}, \varphi(A) < +\infty$ 由于在第一章中我们约定集函数不取 $-\infty$ 为值, 所以这时 φ 是有限的. 记

$$\sup_{A \in \mathscr{A}} \varphi(A) = \sup \varphi.$$

于是存在一集序列 $\{A_n\} \subset \mathscr{A}$, 使

$$\lim_{n \to \infty} \varphi(A_n) = \sup \varphi.$$

令

$$A = \bigcup_{n=1}^{\infty} A_n.$$

按下述方式将 A 逐次加细划分成若干不相交子集的并.

(1) $A = A_1 + (A - A_1)$,

(2) $A = A_1 \cap A_2 + A_1 \cap (A - A_2) + (A - A_1) \cap A_2 + (A - A_1) \cap (A - A_2)$,

$\cdots$,

(n) $A = \sum A_{n,m}, A_{n,m} = A'_1 \cap \cdots \cap A'_n, A'_k = A_k$ 或 $A - A_k$,

$\cdots$.

容易看出, 每一划分都是在前一次划分的基础上加细划分而得来的, 所以, 当 $n' > n$ 时, 每一个 $A_{n,m}$ 都是若干个 $A_{n',m'}$ 的并, 令

$$B_n = \sum_{\varphi(A_{n,m}) \geqslant 0} A_{n,m},$$

即一切使 φ 的值非负的 $A_{n,m}$(n 固定) 的并. 如果这样的 $A_{n,m}$ 不存在, 则令 $B_n = \varnothing$. 为清楚起见, 我们用 $A^+_{n,m}$ 表示使 φ 的值非负的 $A_{n,m}$, 于是 $B_n = \sum A^+_{n,m}$. 根据 $A_{n,m}$ 的取法, 当 $n' > n$ 时, $A^+_{n',m}$ 或是包含在 B_n 中或是与 B_n 不

相交, 所以

$$\begin{aligned}&B_n \cup B_{n+1} \cup \cdots \cup B_{n'}\\ =& B_n + \sum_{A^+_{n+1,m}\cap B_n=\varnothing} A^+_{n+1,m} + \cdots + \sum_{A^+_{n',m}\cap(B_n\cup\cdots\cup B_{n'-1})=\varnothing} A^+_{n',m}.\end{aligned}$$

后一表示式中的集都是两两不交的. 所以利用可加性

$$\begin{aligned}\varphi(B_n\cup\cdots\cup B_{n'}) =& \varphi(B_n) + \sum_{A^+_{n+1,m}\cap B_n=\varnothing} \varphi(A^+_{n+1,m})\\ &+\cdots+ \sum_{A^+_{n',m}\cap(B_n\cup\cdots\cup B_{n'-1})=\varnothing} \varphi(A^+_{n',m}) \geqslant \varphi(B_n),\end{aligned}$$

同时 $A_n = \sum A'_1 \cap \cdots \cap A'_{n-1} \cap A_n$, 可见, A_n 是一些 $A_{n,m}$ 的和, 因而

$$\varphi(A_n) \leqslant \varphi(B_n) \leqslant \varphi(B_n\cup\cdots\cup B_{n'}).$$

由此令 $n' \to \infty$, 利用 σ-可加集函数的下连续性

$$\varphi(A_n) \leqslant \varphi\left(\bigcup_{k=n}^{\infty} B_k\right).$$

令 $C_n = \bigcup\limits_{k=n}^{\infty} B_k$, 易见 $C_n \downarrow C \in \mathscr{A}$. 注意到 φ 是有限的, 利用 φ 的上连续性, 得出

$$\sup\varphi = \lim_{n\to\infty} \varphi(A_n) \leqslant \lim_{n\to\infty} \varphi(C_n) = \varphi(C).$$

但 $C \in \mathscr{A}, \varphi(C) \leqslant \sup\varphi$, 于是就找到一个 $C \in \mathscr{A}$ 使 $\varphi(C) = \sup\varphi$. □

定理 2(Hahn 分解定理) 如果 φ 是 σ-代数 $\mathscr{A}$ 上 σ-可加集函数 (不取 $-\infty$ 为值), 则

i) 存在集 $D \in \mathscr{A}$, 使得对每一集 $A \in \mathscr{A}$ 有

$$\varphi(A\cap D) = \inf_{\substack{B\subset A\\ B\in\mathscr{A}}} \varphi(B), \tag{2}$$

$$\varphi(A\cap D^c) = \sup_{\substack{B\subset A\\ B\in\mathscr{A}}} \varphi(B). \tag{3}$$

ii) 若令

$$\varphi^+(A) = \varphi(A\cap D^c), \tag{4}$$

$$\varphi^-(A) = -\varphi(A\cap D), A \in \mathscr{A}, \tag{5}$$

$$\bar{\varphi}(A) = \varphi^+(A) + \varphi^-(A), \tag{6}$$

则 φ^+, φ^- 和 $\bar{\varphi}$ 都是 $\mathscr{A}$ 上的测度, φ^- 有限, 且

$$\varphi = \varphi^+ - \varphi^-.$$

φ^+, φ^- 和 $\bar{\varphi}$ 分别称为 φ 的**上、下和全变差**.

证　i) 根据定理 1 , 存在 $D \in \mathscr{A}$, 使 $\varphi(D) = \inf\varphi$, 且 $-\infty < \varphi(D) \leqslant 0 = \varphi(\varnothing)$, 今证 D 满足 (2), (3).

要证 (2), 只需证明对于任意 $B \subset A, B \in \mathscr{A}$, 有 $\varphi(A \cap D) \leqslant \varphi(B)$. 用反证法, 如若不然, 即存在 $B \subset A$ 使 $\varphi(A \cap D) > \varphi(B)$, 注意到 $\varphi(A^c \cap D)$ 有限且 $A^c \cap D$ 与 B 不相交, 于是有

$$\begin{aligned}\varphi(D) &= \varphi(A \cap D) + \varphi(A^c \cap D) > \varphi(B) + \varphi(A^c \cap D)\\ &= \varphi(B + A^c \cap D).\end{aligned}$$

这与 D 的取法不合, 这就证明了 (2).

当 $\varphi(A \cap D^c) = \infty$ 时, (3) 式显然成立, 而当 $\varphi(A \cap D^c) < \infty$ 时, 对任意 $B \subset A$,

$$\varphi(A \cap D^c) + \varphi(A \cap D) = \varphi(A) = \varphi(B) + \varphi(B^c \cap A)$$

有限, 同时因为 $B^c \cap A \subset A$, 利用 (2)

$$\varphi(A \cap D) \leqslant \varphi(B^c \cap A),$$

所以

$$\varphi(A \cap D^c) \geqslant \varphi(B).$$

这就证明了 (3).

ii) 由 (2)(3) 的右端看出 φ^+, φ^- 恒取非负值, 由 (2), (3) 的左端可知 φ^+, φ^- (因而 $\bar{\varphi}$) 是 σ-可加集函数, 且 $\varphi = \varphi^+ - \varphi^-$, 而由 $\varphi(D)$ 有限又知 φ^- 有限, 于是定理获证. □

定理 2 中的 φ 是定义在 σ-代数上的, 如果 φ 是定义在集代数上, 则有

定理 3　若 φ 是定义在集代数 $\mathscr{F}$ 上的 σ-可加集函数且 φ 有下界, 则

$$\varphi = \varphi^+ - \varphi^-.$$

其中

$$\varphi^-(A) = -\inf_{B \subset A} \varphi(B), \ A \in \mathscr{F},$$

$$\varphi^+(A) = \varphi(A) + \varphi^-(A), \ A \in \mathscr{F}.$$

且 φ^+, φ^- 都是 $\mathscr{F}$ 上的测度.

证 首先由 φ^- 的定义, 对一切 $A \in \mathscr{F}$,

$$-\varphi^-(A) = \inf_{\substack{B\subset A\\ B\in\mathscr{F}}} \varphi(B) \leqslant \varphi(\varnothing) = 0.$$

所以 $\varphi^-(A) \geqslant 0$.

其次, 由 φ 有下界 (设为 $-K$), 对一切 $A \in \mathscr{F}$.

$$-\varphi^-(A) = \inf_{B\subset A} \varphi(B) \geqslant -K$$

故 $\varphi^-(A) \leqslant K$.

再次, 往证 φ^- 是 σ-可加的. 若 $A, A_n \in \mathscr{F}, n = 1, 2, \cdots, A_n$ 两两不交且 $A = \sum\limits_{n=1}^{\infty} A_n$, 则对一切 $B \subset A, B \in \mathscr{F}$ 有 $B = \sum\limits_{n=1}^{\infty} A_n \cap B$, 显然 $A_n \cap B \in \mathscr{F}$, 由 φ 的 σ-可加性有

$$\varphi(B) = \sum_{n=1}^{\infty} \varphi(A_n \cap B) \geqslant \sum_{n=1}^{\infty} \inf_{\substack{B_n\subset A_n\\ B_n\in\mathscr{F}}} \varphi(B_n) = -\sum_{n=1}^{\infty} \varphi^-(A_n).$$

因而

$$-\varphi^-(A) = \inf_{\substack{B\subset A\\ B\in\mathscr{F}}} \varphi(B) \geqslant -\sum_{n=1}^{\infty} \varphi^-(A_n),$$

即

$$\varphi^-(A) \leqslant \sum_{n=1}^{\infty} \varphi^-(A_n).$$

为了证明相反的不等式成立, 对一切 $\varepsilon > 0$ 取 $C_n \subset A_n, C_n \in \mathscr{F}$, 使

$$\varphi(C_n) \leqslant \inf_{\substack{B_n\subset A_n\\ B_n\in\mathscr{F}}} \varphi(B_n) + \frac{\varepsilon}{2^n} = -\varphi^-(A_n) + \frac{\varepsilon}{2^n}.$$

(这是由于 $0 \geqslant \inf\limits_{\substack{B_n\subset A_n\\ B_n\in\mathscr{F}}} \varphi(B_n) = -\varphi^-(A_n) \geqslant -K$, 故这样的 C_n 存在.)

由于 $A_n, n = 1, 2, \cdots$ 两两不交, 故 $C_n, n = 1, 2, \cdots$ 两两不交, 于是对一切自然数 l,

$$\varphi\left(\sum_{n=1}^{l} C_n\right) = \sum_{n=1}^{l} \varphi(C_n) \leqslant -\sum_{n=1}^{l} \varphi^-(A_n) + \varepsilon.$$

又因为 $\sum\limits_{n=1}^{l} C_n \subset \sum\limits_{n=1}^{l} A_n \subset A$, 且 $\sum\limits_{n=1}^{l} C_n \in \mathscr{F}$, 所以

$$\varphi\left(\sum_{n=1}^{l} C_n\right) \geqslant \inf_{\substack{B\subset A\\ B\in\mathscr{F}}} \varphi(B) = -\varphi^-(A)$$

把这两个不等式结合起来有

$$-\varphi^-(A) \leqslant -\sum_{n=1}^{l} \varphi^-(A_n) + \varepsilon,$$

即

$$\varphi^-(A) \geqslant \sum_{n=1}^{l} \varphi^-(A_n) - \varepsilon.$$

先令 $l \to \infty$, 再令 $\varepsilon \to 0$. 即得

$$\varphi^-(A) \geqslant \sum_{n=1}^{\infty} \varphi^-(A_n).$$

于是

$$\varphi^-(A) = \sum_{n=1}^{\infty} \varphi^-(A_n).$$

φ^- 是 $\mathscr{F}$ 上的测度获证.

最后我们证明 φ^+ 也是 $\mathscr{F}$ 上的测度. 首先由于 $\varphi^-(A)$ 有界, 故 $\varphi^+(A) = \varphi(A) + \varphi^-(A)$ 有定义. 其次, 由于 $-\varphi^-(A) = \inf\limits_{B \subset A} \varphi(B) \leqslant \varphi(A)$, 故 $\varphi(A) + \varphi^-(A) \geqslant 0$, 因而 φ^+ 非负, 再由 φ, φ^- 都是 $\mathscr{F}$ 上 σ-可加集函数且 φ^- 有界, 因而 φ^+ 也是 σ-可加集函数, 即 φ^+ 也是 $\mathscr{F}$ 上的测度. 最后由 φ^- 有界得出

$$\varphi(A) = \varphi^+(A) - \varphi^-(A), \quad A \in \mathscr{F}.$$

定理证毕. □

对于 φ 有上界的情形, 可以有类似的结论, 就不在此赘述了.

定理 2 和定理 3 的意义在于指出了在集代数 $\mathscr{F}$ 上定义的 σ-可加集函数都可以表为两个测度的差 (因而称为**符号测度**), 从而在许多情况下, 关于一般 σ-可加集函数的研究, 可以归结到关于测度的研究, 而测度比一般 σ-可加集函数更容易掌握些.

3.7.2 不定积分与 Lebesgue 分解定理

在 § 3 我们给出了在测度空间 $(\Omega, \mathscr{A}, \mu)$ 上积分存在的可测函数 f 的不定积分的定义, 并且证明了 f 的不定积分

$$\varphi(A) = \int_A f d\mu$$

是 $\mathscr{A}$ 上 σ-可加集函数. 这一小节我们要解决本节的中心问题, 即在什么条件下集函数 φ 可以表示成某一可测函数关于 μ 的不定积分.

容易看出, 若 $\int f^- d\mu < \infty$, 则 $\int f d\mu$ 存在, 因而 f 的不定积分一定存在, 如

无特别声明, 我们约定 f 满足此条件, 因而不定积分 φ 不取 $-\infty$ 为值.

定理 4 不定积分 $\varphi(A)=\int_A fd\mu, A\in\mathscr{A}$, 具有下列性质:

a) 若 $\mu(A)=0$, 则 $\varphi(A)=0$;

b) φ 是 $\mathscr{A}$ 上 σ-可加集函数;

c) 若 f a.e. 有限, μ 是 σ-有限测度, 则 φ 是 σ-有限集函数. 特别, 当 f 可积时, φ 有限.

证 a), b) 显然成立. 今往证 (c) 成立.

由于 μ 是 σ-有限测度, 存在 $\{A_n\}\subset\mathscr{A}, A_n$ 两两不交, $\mu(A_n)<\infty, n=1,2,\cdots$, 且使 $\Omega=\sum\limits_{n=1}^{\infty}A_n$; 另一方面, 若令 $B=\{\omega: f(\omega)=\pm\infty\}, B_m=\{\omega: m\leqslant f(\omega)<m+1\}, m=\cdots,-1,0,1,\cdots$, 则 $\Omega=B+\sum\limits_{m=-\infty}^{\infty}B_m$, 由假设知 $\mu(B)=0$, 因而 $\varphi(B)=0$, 又因为

$$\Omega=B+\sum_{m=-\infty}^{\infty}\sum_{n=1}^{\infty}(A_n\cap B_m),$$

而

$$-\infty<m\mu(A_n\cap B_m)\leqslant\varphi(A_n\cap B_m)=\int_{A_n\cap B_m}fd\mu\leqslant(m+1)\mu(A_n\cap B_m)<+\infty,$$

由此可见 φ 是 σ-有限集函数. □

不定积分的性质 a) 称为 μ-连续性. 一般地, 我们有下面的

定义 1 设 $(\Omega,\mathscr{A},\mu)$ 是一测度空间, φ 是 $\mathscr{A}$ 上的一个集函数, 如果对任意的 $A\in\mathscr{A}$ 且 $\mu(A)=0, \varphi(A)=0$ 成立, 则称 φ 是 **μ-连续的**. 记作 $\varphi\ll\mu$.

为了证明定理 4 的逆定理, 我们先证明更广一点的 Lebesgue 分解定理, 即证明任何一个 σ-可加集函数可分解为两部分, 一部分是不定积分, 因而是 μ-连续的, 另一部分是所谓 μ-奇异的. 为此, 先给出

定义 2 设 $(\Omega,\mathscr{A},\mu)$ 是一测度空间, φ 是 $\mathscr{A}$ 上的一个集函数, 如果存在 $N\in\mathscr{A}, \mu(N)=0$, 使得对于任意的 $A\in\mathscr{A}$, 都有

$$\varphi(A\cap N^c)=0,$$

其中 $N^c=\Omega-N$, 则称 φ 是 **μ-奇异的**.

定理 5(Lebesgue 分解定理) 设 $(\Omega,\mathscr{A},\mu)$ 是一测度空间, φ 是 $\mathscr{A}$ 上的 σ-可加集函数, 且 μ 和 φ 都是 σ-有限的, 则 $\varphi=\varphi_c+\varphi_s$, 其中 φ_c 是 $(\Omega,\mathscr{A},\mu)$ 上某一

有限可测函数 f 的不定积分, 因而具有 σ-可加性及 μ-连续性, 而 φ_s 是 μ-奇异的 σ-可加集函数. 若 $\varphi > -\infty$, 则 $\varphi_c > -\infty, \varphi_s > -\infty$. 并且这样的分解是唯一的, 同时 f 也由 φ a.e. 唯一确定, 即若 f_1 满足上述要求, 则 $f_1 = f$, a.e..

此分解称为 φ 的 Lebesgue 分解.

证　i) 分解的唯一性: 设有两种分解,

$$\varphi = \varphi_c + \varphi_s = {\varphi'}_c + {\varphi'}_s,$$

其中 $\varphi_c, {\varphi'}_c$ 是 σ-可加、μ-连续的, 而 $\varphi_s, {\varphi'}_s$ 是 σ-可加、μ-奇异的. 因此存在 $N_1, N_2 \in \mathscr{A}, \mu(N_i) = 0, i = 1, 2$, 使得对任何 $A \in \mathscr{A}$,

$$\varphi_s(A \cap N_1^c) = {\varphi'}_s(A \cap N_2^c) = 0,$$

令 $N = N_1 \cup N_2$, 则 $N^c = N_1^c \cap N_2^c, \mu(N) = 0$. 今往证对任何 $A \in \mathscr{A}$,

$$\varphi_c(A) = {\varphi'}_c(A), \quad \varphi_s(A) = {\varphi'}_s(A).$$

先证明 $\varphi(A)$ 有限的情形. 由假设

$$\varphi(A) = \varphi_c(A) + \varphi_s(A) = {\varphi'}_c(A) + {\varphi'}_s(A),$$

由于 $\varphi(A)$ 有限, 因此 ${\varphi'}_c(A), \varphi_s(A)$ 有限, 于是

$$\varphi_c(A) - \varphi'_c(A) = \varphi'_s(A) - \varphi_s(A). \tag{7}$$

再由 $\varphi(A)$ 有限知 $\varphi(A \cap N^c)$ 有限, 同样的论证有

$$\varphi_c(A \cap N^c) - \varphi'_c(A \cap N^c) = \varphi'_s(A \cap N^c) - \varphi_s(A \cap N^c). \tag{8}$$

利用 φ_c, φ'_c 的 μ 连续性, φ_s, φ'_s 的 μ-奇异性及它们的可加性, 即得

$$\begin{aligned}\varphi_c(A) - \varphi'_c(A) &= \varphi_c(A \cap N^c) - \varphi'_c(A \cap N^c) \\ &= \varphi'_s(A \cap N^c) - \varphi_s(A \cap N^c) \\ &= \varphi'_s(A \cap N_1^c \cap N_2^c) - \varphi_s(A \cap N_2^c \cap N_1^c) = 0,\end{aligned}$$

再利用 (7) 即得

$$\varphi'_s(A) - \varphi_s(A) = 0.$$

再证 $\varphi(A) = \infty$ 的情形. 由于 φ 是 σ-有限的, 故存在 $\{A_n\} \subset \mathscr{A}, A_n$ 两两不交使 $\Omega = \sum_{n=1}^{\infty} A_n, \varphi(A_n)$ 有限, 于是由 σ-可加性及 i) 知

$$\varphi_c(A) = \sum_n \varphi_c(A \cap A_n) = \sum_n \varphi'_c(A \cap A_n) = \varphi'_c(A),$$

$$\varphi_s(A) = \sum_n \varphi_s(A \cap A_n) = \sum_n \varphi_s'(A \cap A_n) = \varphi_s'(A).$$

于是分解的唯一性得证.

ii) f 的唯一性: 设 f 和 f_1 均满足定理的要求, 则由它们产生的 φ 的 μ-连续部分

$$\varphi_c(A) = \int_A f d\mu, \quad \varphi_c'(A) = \int_A f_1 d\mu,$$

根据分解的唯一性, 是相等的. 即对任意 $A \in \mathscr{A}$,

$$\int_A f d\mu = \int_A f_1 d\mu.$$

由积分的性质 (§ 2.2 定理 2, 2)a)) 知 $f = f_1$, a.e.. 这就证明了 f 由 φ a.e. 唯一决定.

iii) 分解的存在性: 由 Hahn 分解定理知任何一个 σ-可加集函数均可表示成两个测度的差, 而有限的情形又是 σ-有限的基础, 因而我们分以下三个步骤证明分解的存在性.

1) μ, φ 是有限测度;

2) μ, φ 是 σ-有限测度;

3) φ 是 φ-有限、σ-可加集函数, μ 是 σ-有限测度.

先证 1) 设

$$\Phi = \left\{ f : f \geqslant 0, \text{可测}, \int_A f d\mu \leqslant \varphi(A), \text{对一切} A \in \mathscr{A} \right\}.$$

因为 $f = 0 \in \Phi$, 故 Φ 非空. 取一函数序列 $\{f_n\} \subset \Phi$ 使

$$\lim_{n \to \infty} \int f_n d\mu = \sup_{f \in \Phi} \int f d\mu = \alpha \leqslant \varphi(\Omega) < \infty,$$

令 $g_n = \sup\limits_{k \leqslant n} f_k$, 则 $0 \leqslant g_n \uparrow f = \sup\limits_{n \geqslant 1} f_n$, 又设

$$A_k = \{\omega : g_n(\omega) = f_k(\omega)\},$$

$$B_1 = A_1, B_k = A_1^c \cap \cdots \cap A_{k-1}^c \cap A_k, k = 2, \cdots, n,$$

则 $B_k, k = 1, \cdots, n$ 两两不交且

$$\sum_{k=1}^n B_k = \bigcup_{k=1}^n A_k = \Omega.$$

对任一 $A \in \mathscr{A}$,

$$\int_A g_n d\mu = \sum_{k=1}^n \int_{A\cap B_k} f_k d\mu \leqslant \sum_{k=1}^n \varphi(A\cap B_k) = \varphi(A).$$

令 $n\to\infty$, 利用单调收敛定理得到

$$\int_A f d\mu \leqslant \varphi(A), \int f d\mu = \alpha.$$

故知 $f \in \Phi$. 今往证

$$0 \leqslant \varphi_c(A) \triangleq \int_A f d\mu, A\in\mathscr{A},$$
$$0 \leqslant \varphi_s(A) \triangleq \varphi(A) - \varphi_c(A), A\in\mathscr{A},$$

即是定理所要求的分解，为此只需证 φ_s 是 μ-奇异的.

令 $\varphi_n = \varphi_s - \frac{1}{n}\mu$, 显然对一切 $n\geqslant 1, \varphi_n$ 是有限 σ-可加集函数, 利用 Hahn 分解定理, 存在 $D_n \in \mathscr{A}$ 使对一切 $A\in\mathscr{A}$,

$$\varphi_n(A\cap D_n) \leqslant 0, \varphi_n(A\cap D_n^c) \geqslant 0.$$

取 $D = \bigcap\limits_{n=1}^\infty D_n$, 则对任意 $A\in\mathscr{A}$ 及 $n\geqslant 1$,

$$\begin{aligned}\varphi_s(A\cap D) - \frac{1}{n}\mu(A\cap D) &= \varphi_n(A\cap D)\\ &= \varphi_n(A\cap D\cap D_n) \leqslant 0,\end{aligned}$$

即

$$0 \leqslant \varphi_s(A\cap D) \leqslant \frac{1}{n}\mu(A\cap D),$$

令 $n\to\infty$, 注意到 μ 是有限的, 对任意 $A\in\mathscr{A}$ 有

$$\varphi_s(A\cap D) = 0.$$

于是为了证明 φ_s 的 μ-奇异性, 只需再证明

$$\mu(D^c) = 0.$$

事实上,

$$\begin{aligned}\int_A \left(f + \frac{1}{n}\chi_{D_n^c}\right) d\mu &= \varphi_c(A) + \frac{1}{n}\mu(A\cap D_n^c)\\ &= \varphi(A) - \varphi_s(A) + \frac{1}{n}\mu(A\cap D_n^c).\end{aligned}$$

注意到 $\varphi_s(A\cap D)=0$ 及 $D^c\supset D_n^c$,

$$\begin{aligned}\text{上式}&=\varphi(A)-\varphi_s(A\cap D^c)+\frac{1}{n}\mu(A\cap D_n^c)\\&\leqslant\varphi(A)-\varphi_s(A\cap D_n^c)+\frac{1}{n}\mu(A\cap D_n^c)\\&=\varphi(A)-\varphi_n(A\cap D_n^c)\leqslant\varphi(A).\end{aligned}$$

这就是说, $f+\frac{1}{n}\chi_{D_n^c}\in\varPhi$, 因此

$$\begin{aligned}\alpha\geqslant\int_\Omega\left(f+\frac{1}{n}\chi_{D_n^c}\right)d\mu&=\varphi_c(\Omega)+\frac{1}{n}\mu(D_n^c)\\&=\alpha+\frac{1}{n}\mu(D_n^c).\end{aligned}$$

由于 $0\leqslant\alpha<\infty$, 故知 $\mu(D_n^c)=0$, 而

$$D^c=\bigcup_{n=1}^\infty D_n^c,$$

所以

$$\mu(D^c)=0.$$

这就证明了 φ_s 是 μ-奇异的. 于是在 μ,φ 是有限测度的情形, 分解的存在性获证.

2) 现在证明 μ,φ 是 σ-有限测度时分解的存在性: 我们将 Ω 划分为 $\Omega=\sum\limits_{k=1}^\infty A_k, A_k$ 两两不交且使 $\varphi(A_k),\mu(A_k)$ 有限, 设 $\mathscr{A}\cap A_k=\{A_k\cap A:A\in\mathscr{A}\}$, 显然它是以 A_k 为空间的 σ-代数, 且在此 σ-代数上, μ 和 φ 都是有限测度, 于是利用 1) 中的结论, 对每一 k,

$$\varphi(A_k\cap A)=\varphi_c^{(k)}(A_k\cap A)+\varphi_s^{(k)}(A_k\cap A), A\in\mathscr{A},$$

其中

$$0\leqslant\varphi_c^{(k)}(A_k\cap A)=\int_{A_k\cap A}f_k(\omega)d\mu,$$

$f_k\geqslant 0$ 且有限, $\mathscr{A}\cap A_k$ 可测,

$$0\leqslant\varphi_s^{(k)}(A_k\cap A\cap N_k^c)=0, \mu(N_k)=0, N_k\in\mathscr{A}\cap A_k.$$

令

$$\varphi_c(A)=\sum_{k=1}^\infty\varphi_c^{(k)}(A_k\cap A),$$

$$\varphi_s(A)=\sum_{k=1}^\infty\varphi_s^{(k)}(A_k\cap A).$$

由于 $\varphi_c^{(k)},\varphi_s^{(k)}$ 都是非负的, 故 φ_c,φ_s 有意义, 这样对任意 $A\in\mathscr{A}$,

$$\varphi(A)=\sum_{k=1}^{\infty}\varphi(A_k\cap A)=\sum_{k=1}^{\infty}\varphi_c^{(k)}(A_k\cap A)+\sum_{k=1}^{\infty}\varphi_s^{(k)}(A_k\cap A)$$
$$=\varphi_c(A)+\varphi_s(A).$$

其次, φ_c 是函数 $f=\sum\limits_{k=1}^{\infty}f_k\chi_{A_k}$ 的不定积分. f 是 $\mathscr{A}$ 可测的 (请读者自证). 只要利用单调收敛定理 (注意 $f_k\geqslant 0$) 即可看出

$$\varphi_c(A)=\sum_{k=1}^{\infty}\int_{A_k\cap A}f_k d\mu=\sum_{k=1}^{\infty}\int_A f_k\chi_{A_k}d\mu$$
$$=\int_A\sum_{k=1}^{\infty}f_k\chi_{A_k}d\mu=\int_A f d\mu.$$

最后, 我们证明 φ_s 是 μ-奇异的. 取 $N=\bigcup\limits_{k=1}^{\infty}N_k\in\mathscr{A}$, 这时 $N^c=\bigcap\limits_{k=1}^{\infty}N_k^c$, 显然 $\mu(N)=0$. 而对于任意 $A\in\mathscr{A}$,

$$\varphi_s(A\cap N^c)=\sum_{k=1}^{\infty}\varphi_s^{(k)}(A_k\cap A\cap N^c)$$
$$=\sum_{k=1}^{\infty}\varphi_s^{(k)}(A_k\cap A\cap N^c\cap N_k^c)=0.$$

这样就证明了当 μ 和 φ 都是 σ-有限测度时分解的存在性.

我们还注意到当 φ 是测度时, 分解的两部分都是测度, 同时 μ 连续部分是一非负可测函数的不定积分.

3) 最后证明当 φ 是任意 σ-有限、σ-可加集函数, μ 是 σ-有限测度时分解的存在性.

利用 Hahn 分解定理, 可知

$$\varphi=\varphi^+-\varphi^-,$$

其中 φ^+ 和 φ^- 是两个测度, φ^- 有限, 利用 2) 中结论,

$$\varphi^+=\varphi_c^++\varphi_s^+,$$
$$\varphi^-=\varphi_c^-+\varphi_s^-,$$

其中

$$\varphi_c^+(A)=\int_A f_1(\omega)d\mu,\varphi_s^+(A\cap N_1^c)=0,\mu(N_1)=0,$$

$$\varphi_c^-(A)=\int_A f_2(\omega)d\mu, \varphi_s^-(A\cap N_2^c)=0, \mu(N_2)=0$$

对任意 $A\in\mathscr{A}$ 成立. 令

$$\varphi_c=\varphi_c^+-\varphi_c^-,\quad \varphi_s=\varphi_s^+-\varphi_s^-,$$

则

$$\varphi=\varphi_c+\varphi_s,$$

并且 φ_c 是 $f\triangleq f_1-f_2$ 的不定积分 (注意 f_2 可积, 因而 f_1-f_2 积分存在). 即对任意 $A\in\mathscr{A}$,

$$\varphi_c(A)=\varphi_c^+(A)-\varphi_c^-(A)=\int_A f_1d\mu-\int_A f_2d\mu=\int_A fd\mu.$$

而 φ_s 是 μ-奇异的. 取 $N=N_1\cup N_2$, 则 $\mu(N)=0, N^c=N_1^c\cap N_2^c$, 对任意 $A\in\mathscr{A}$,

$$\varphi_s(A\cap N^c)=\varphi_s^+(A\cap N^c)-\varphi_s^-(A\cap N^c)$$

$$=\varphi_s^+(A\cap N_2^c\cap N_1^c)-\varphi_s^-(A\cap N_1^c\cap N_2^c)=0.$$

定理到此全部证毕. □

在定理 5 中, 特别当 φ 具有 μ-连续性时, 由分解的唯一性, $\varphi=\varphi_c, \varphi_s=0$, 于是得到了下面重要的定理.

定理 6(Radon-Nikodym 定理) 设 μ 是 Ω 中 σ-代数 $\mathscr{A}$ 上的 σ-有限测度, 如果 $\mathscr{A}$ 上的集函数 φ 是 σ-有限、σ-可加且 μ-连续的, 则 φ 是测度空间 $(\Omega,\mathscr{A},\mu)$ 上某一有限可测函数 f 的不定积分, 且 f 由 φ 几乎唯一确定.

定理 7(Radon-Nikodym 定理的推广) 设 μ 是 Ω 是 σ-代数 $\mathscr{A}$ 上的 σ-有限测度, φ 是 $\mathscr{A}$ 上的 μ-连续的 σ-可加集函数 (不一定 σ-有限), 则 φ 是测度空间 $(\Omega,\mathscr{A},\mu)$ 上某一可测函数 f (不一定 a.e. 有限) 的不定积分, 且 f 由 φ 几乎唯一确定.

证 我们只证 μ 是有限测度及 φ 是测度的情形, 其他情形由 Hahn 分解定理及与定理 5 相同的手法容易得到. 请读者完成.

设

$$\mathscr{B}=\left\{A:\varphi\text{在}A\text{上}\sigma-\text{有限, 即存在}\{A_n\}\subset\mathscr{A},\text{对一切}n,\varphi(A_n)<\infty,\text{使}\bigcup_n A_n=A\right\}.$$

令

$$s=\sup_{B\in\mathscr{B}}\mu(B).$$

则有一集序列 $\{B_n\} \subset \mathscr{B}$, 使

$$s = \lim_{n\to\infty} \mu(B_n).$$

令

$$B = \bigcup_{n=1}^{\infty} B_n.$$

因每一 B_n 关于 φ 是 σ-有限的, 故 B 也 σ-有限, 因而 $B \in \mathscr{B}.s = \mu(B)$.

集类 $\mathscr{A}_1 \triangleq B \cap \mathscr{A}, \mathscr{A}_2 \triangleq B^c \cap \mathscr{A}$ 分别是以 B 和 B^c 为空间的 σ-代数, φ 在 $\mathscr{A}_1$ 及 $\mathscr{A}_2$ 上的限制也分别是 $\mathscr{A}_1$ 及 $\mathscr{A}_2$ 上的 μ-连续测度, 且由 B 的定义和 φ 在 $\mathscr{A}_1$ 上是 σ-有限的, 故由 Radon-Nikodym 定理知有一定义在 B 上且 $\mathscr{A}_1$ 可测的函数 f_1 存在, 使得

$$\varphi(A_1) = \int_{A_1} f_1(\omega) d\mu, \quad A_1 \in \mathscr{A}_1. \tag{9}$$

令 $f_2(\omega) = \infty, \omega \in B^c$, 今往证

$$\varphi(A_2) = \int_{A_2} f_2(\omega) d\mu, \quad A_2 \in \mathscr{A}_2. \tag{10}$$

为此只需证在 $\mathscr{A}_2$ 上若 $\mu(A_2) > 0$, 则 $\varphi(A_2) = \infty$.(若 $\mu(A_2) = 0$, 则 $\varphi(A_2) = 0$, 由 φ 是 μ-连续的显然成立.) 事实上若存在 $A_2 \in \mathscr{A}_2, \mu(A_2) > 0$, 而 $\varphi(A_2) < \infty$, 由于 $A_2 \cap B = \varnothing$, 故 $\mu(A_2 + B) = \mu(A_2) + \mu(B) > s$, 而 $A_2 + B$ 在 $\mathscr{A}$ 上是 σ-有限的, 故 $A_2 + B \in \mathscr{B}, \mu(A_2 + B) \leqslant s$. 这一矛盾就证明了若 $A_2 \in \mathscr{A}_2, \mu(A_2) > 0$ 则 $\varphi(A_2) = \infty$, 于是 (10) 成立.

令

$$f(\omega) = \begin{cases} f_1(\omega), & \omega \in B, \\ f_2(\omega) = \infty, & \omega \in B^c. \end{cases}$$

则对任何 $A \in \mathscr{A}$,

$$\varphi(A) = \int_A f d\mu, \quad A \in \mathscr{A}.$$

f 的 a.e. 唯一性由积分的性质 (§2.2 定理 2, 2)a)) 推出.

定义 3　设 $(\Omega, \mathscr{A}, \mu)$ 是 σ-有限测度空间, φ 是 $\mathscr{A}$ 上 μ 连续的 σ-可加集函数, 按定理 7 a.e. 唯一确定的函数 f 称为 φ 关于 μ 的 Radon **导数**, 记作 $\dfrac{d\varphi}{d\mu}$.

由定理 7 及 § 5.3 定理 7 可得

推论　设 λ, μ 是 Ω 中 σ-代数 $\mathscr{A}$ 上的测度, μ 是 σ-有限的, λ 是 μ 连续的. 若 f 是 $(\Omega, \mathscr{A})$ 上的可测函数, 则 $\int f d\lambda$ 存在当且仅当 $\int f \dfrac{d\lambda}{d\mu} d\mu$ 存在, 且在积分存在时, 对任意 $A \in \mathscr{A}$

$$\int_A f d\lambda = \int_A f \frac{d\lambda}{d\mu} d\mu.$$

3.7.3 分布函数的分解

定理 8 任一有界的分布函数 $F(x_1,\cdots,x_n)$ 都可以分解为三个分布函数的和, 即

$$F=F_c+F_d+F_s,$$

其中 F_c 是 $R^{(n)}$ 上非负有限 Borel 可测函数 p 关于 L 测度的不定积分, 即对一切 $(x_1,\cdots,x_n)\in R^{(n)}$

$$F_c(x_1,\cdots,x_n)=\int_{-\infty}^{x_1}\cdots\int_{\infty}^{x_n}p(t_1,\cdots,t_n)dt_1\cdots dt_n,$$

F_c 称为F **的绝对连续部分**; F_d 是一 n 维阶梯形函数, 即存在可数个点 $(a_{1k_1},\cdots,a_{nk_n})$ 及 $p_{k_1,\cdots,k_n}\geqslant 0, k_1,k_2,\cdots,k_n=1,2,\cdots$ 使

$$F_d(x_1,\cdots,x_n)=\sum_{\substack{a_{ik_i}<x_i\\ i=1,\cdots,n}}p_{k_1}\cdots k_n.$$

F_d 称为 F 的**离散部分**; F_s 是一分布函数, 它所对应的 L–S 测度关于 L 测度是奇异的, 并且 F_s 的差分还是右连续的, 即 $\lim\limits_{b\to a+}\Delta_{b,a}F=0$, 其中 $a=(a_1,\cdots,a_n)$ 是 $R^{(n)}$ 中的任一点, $b=(b_1,\cdots,b_n)\in R^{(n)}, b\to a+$ 表示对每一 $k, b_k\to a_k+$.F_s 称为 F 的**奇异部分**.

证 设 μ 是对应于 F 的 L–S 测度, 根据 Lebesgue 分解定理 (μ 看作那里的 φ, Lebesgue 测度 Λ 看作那里的 μ), 存在分解式

$$\mu=\mu_c+\mu_s',$$

其中 μ_c 是 Λ-连续的, 可表示成非负有限 Borel 可测函数的不定积分,

$$\mu_c(A)=\int_A p(t_1,\cdots,t_n)dt_1\cdots dt_n,$$

并存在一个集合 $N, \Lambda(N)=0$, 且对任一 Borel 可测集 (可以是 Lebesgue 可测集)A,

$$\mu_s'(A\cap N^c)=0. \tag{11}$$

令

$$\begin{aligned}F_c(x_1,\cdots,x_n)&=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}p(t_1,\cdots,t_n)dt_1\cdots dt_n\\ &=\mu_c(((-\infty,\cdots,-\infty),(x_1,\cdots,x_n))),\end{aligned}$$

就得出 F 的绝对连续部分, 在此, 由于 F 有界, 因而 μ_c,μ_s' 都是有限测度, 从而 F_c 是有界分布函数.

为了得出后两部分, 将 μ_s' 再进行分解. 为此先证明使 $\mu_s'(\{(a_1,\cdots,a_n)\})\neq 0$ 的点 $(a_1,\cdots,a_n)$ 至多有可数个. 首先可以证明, 使 $\mu_s'>\dfrac{1}{m}$ 的点只有有限多个, 如若不然, 在 $R^{(n)}$ 中存在无限多个不同的点 $a^{(1)},\cdots,a^{(k)},\cdots$ 使 $\mu_s'(a^{(k)})>\dfrac{1}{m},k=1,2,\cdots$, 则

$$\infty=\sum_{k=1}^{\infty}\mu_s'(\{a^{(k)}\})=\mu_s'(\{a^{(1)},a^{(2)},\cdots\})$$

$$\leqslant\mu_s'(R^{(n)})\leqslant\mu(R^{(n)})<\infty,$$

引出矛盾. 由于使 $\mu_s'\neq 0$ 的点是使 $\mu_s'>\dfrac{1}{m},m=1,2,\cdots$ 的总和, 因而使 $\mu_s'\neq 0$ 的点至多可数.

在一切使 $\mu_s'>0$ 的点中, 每一坐标也至多出现可数多个不同的值, 使各个坐标的不同值依次为

$$\begin{gathered}a_{11},a_{12},\cdots,a_{1k},\cdots\\ a_{21},a_{22},\cdots,a_{2k},\cdots\\ \cdots\\ a_{n1},a_{n2},\cdots,a_{nk},\cdots\end{gathered}$$

则 $(a_{1k_1},\cdots,a_{nk_n}),k_i=1,2,\cdots,i=1,\cdots,n$ 包括了一切使 $\mu_s'>0$ 的点. 设

$$\mu_s'(\{(a_{1_{k_1}},\cdots,a_{n_{k_n}})\})=p_{k_1},\cdots_{k_n},$$

则 $p_{k_1}\cdots k_n\geqslant 0,k_i=1,2,\cdots,i=1,\cdots,n.$

令

$$\mu_d(B)=\sum_{(a_{1k_1},\cdots,a_{nk_n})\in B}\mu_s'(\{(a_{1k_1},\cdots,a_{n_{kn}})\})$$

$$=\sum_{(a_{1k_1},\cdots,a_{nk_n})\in B}p_{k_1\cdots k_n},\qquad B\in\mathscr{B}^{(n)}.$$

如果 B 不含有任何的 $(a_{1k_1},\cdots,a_{nk_n})$, 则 $\mu_d(B)$ 理解为零. 又令

$$\mu_s(B)=\mu_s'(B)-\mu_d(B),\qquad B\in\mathscr{B}^{(n)}.$$

显然 $\mu_s'(B)\geqslant\mu_d(B),\mu_d$ 是测度. 因而 μ_s 也是测度, 且有

$$\mu=\mu_c+\mu_d+\mu_s.$$

令

$$F_d(x_1,\cdots,x_n)=\mu_d(((-\infty,\cdots,-\infty),(x_1,\cdots,x_n)))$$
$$=\sum_{\substack{a_{ik_i}<x_i\\ i=1,\cdots,n}} p_{k_1\cdots k_n},$$
$$F_s=F-F_c-F_d.$$

F_d 是 μ_d 的分布函数, 从而得知

$$\Delta_{b,a}F_s=\Delta_{b,a}F-\Delta_{b,a}F_c-\Delta_{b,a}F_d$$
$$=\mu([a,b))-\mu_c([a,b))-\mu_d([a,b))$$
$$=\mu_s([a,b)).$$

所以 F_s 是 μ_s 的分布函数, 于是已经把 F 表成了三个分布函数之和, 即

$$F=F_c+F_d+F_s.$$

其中 F_c 和 F_d 显然满足定理的要求. 只剩下证明 F_s 对应的测度 μ_s 关于 L 测度 Λ 是奇异的且 F_s 的 n 阶差分右连续.

事实上, 取 (11) 中的集 $N,\Lambda(N)=0$,

$$0\leqslant\mu_s(A\cap N^c)=\mu_s'(A\cap N^c)-\mu_d(A\cap N^c)$$
$$\leqslant\mu_s'(A\cap N^c)=0,\qquad A\in\mathscr{B}^{(n)},$$

故 μ_s 是 Λ-奇异的.

设 $a=(a_1,\cdots,a_n),b=(b_1,\cdots,b_n),a_k<b_k$,

则

$$\Delta_{b,a}F_s=\mu_s([(a_1,\cdots,a_n),(b_1,\cdots,b_n))),\ \lim_{b\to a+}\Delta_{b,a}F_s=\mu_s(\{(a_1,\cdots,a_n)\})$$
$$=\mu_s'(\{(a_1,\cdots,a_n)\})-\mu_d(\{(a_1,\cdots,a_n)\})=0.$$

至此定理证完. □

这个定理表明分布函数的性质有可能从各个分解出的函数的性质推出. 其中绝对连续部分 (如果不恒为零)F_c 如果用 $c=F_c(+\infty,\cdots,+\infty)$ 除它, 则 $\dfrac{1}{c}F_c(x_1,\cdots,x_n)$ 是某一具有分布密度 $\dfrac{1}{c}p(x_1,\cdots,x_n)$ 的连续型随机变量的分布函数, 而离散部分 F_d(如果不恒为零) 除以 $d=F_d(+\infty,\cdots,+\infty)$, 则

$$\frac{1}{d}F_d(x_1,\cdots,x_n)=\sum_{\substack{a_{ik_i}<x_i\\ i=1,\cdots,n}}\frac{1}{d}p_{k_1\cdots k_n}$$

是可能值为 $(a_{1k_1},\cdots,a_{nk_n}),k_1,\cdots,k_n=1,2,\cdots$ 取各值的概率 $P(\xi_1=a_{1k_1},\cdots,\xi_n=a_{nk_n})=\frac{1}{d}p_{k_1\cdots kn}$ 的离散型随机变量的分布函数. 至于奇异部分 F_s 除以常数 $F_s(+\infty,\cdots,+\infty)$ (如果不为零) 后, 也相当于某一随机变量的分布函数. 这个随机变量几乎在所有的地方取值的概率都等于零.(因 $\mu_s(A\cap N^c)=0,\Lambda(N)=0$); 只在很小的范围内 (在 L 测度为零的集 N 内) 取值才不为零. 而取每一点的概率为零. 这样的随机变量在实际中是很少见的. 因而, 离散型和连续型随机变量是两种基本类型的随机变量.

习题及补充

在下面各题中 φ 表示 $(\Omega,\mathscr{A})$ 上 σ-可加集函数, $\varphi=\varphi^+-\varphi^-$ 表示 Hahn 分解, $\bar{\varphi}\triangleq\varphi^++\varphi^-$, μ 表示测度, $A,B,\cdots$ 表示 $\mathscr{A}$ 可测集.

1. 若 $\varphi=\mu_1-\mu_2$, 则必有 $\varphi^+\leqslant\mu_1,\varphi^-\leqslant\mu_2$ (此谓之 Hahn 分解的最小性).

2. 如果 φ 是有限可加的, μ 是有限的, 并且 $\mu(A_n)\to 0$ 时 $\varphi(A_n)\to 0$, 则 φ 是 σ-可加的.

3. 若 $\mu(A_n)\to 0$ 蕴含 $\varphi(A_n)\to 0(\bar{\varphi}(A_n)\to 0)$, 则 φ 是 μ 连续的; 如果 φ 是有限的, 反之亦真. (提示: 若其逆不真, 则存在 $\varepsilon>0$ 与序列 $\{A_n\}$, 使得$\mu(A_n)<\frac{1}{2^n}$ 且 $\bar{\varphi}(A_n)\geqslant\varepsilon$, 令 $B=\varlimsup\limits_{n\to\infty}A_n$, 使得 $\mu(B)=0$ 而 $\bar{\varphi}(B)\geqslant\varepsilon$.)

若 φ 是 σ-有限时如何?

4. 如果 $\mu_j,j=1,2,\cdots$ 是有限测度, 则必存在一个 μ, 使得所有的 μ_j 都是 μ 连续的.$\left(\text{提示: 取 }\mu=\sum\limits_{j=1}^{\infty}\frac{\mu_j}{2^j\mu_j(\Omega)}.\right)$

μ_j 可否换为 φ_j?

5. 微分的形式法则对于 Radon-Nikodym 导数是适用的, 即设 μ,ν 是 $\mathscr{A}$ 上的有限测度, φ,φ' 是 $\mathscr{A}$ 上的 σ-有限的 σ-可加集函数, 设 φ 是 ν 连续的, 并且 ν,φ,φ' 是 μ 连续的, 则

$$\frac{d(\varphi+\varphi')}{d\mu}=\frac{d\varphi}{d\mu}+\frac{d\varphi'}{d\mu},\quad \mu-\text{a.e.}$$

$$\frac{d\varphi}{d\mu}=\frac{d\varphi}{d\nu}\frac{d\nu}{d\mu},\quad \mu-\text{a.e.}$$

6. 设 $\bar{\mu}_n=\sum\limits_{k=1}^{n}\mu_k\to\bar{\mu}$ 并且 $\bar{\nu}_n=\sum\limits_{k=1}^{n}\nu_k\to\bar{\nu}$, 其中带有附标的 μ,ν 皆是 $(\Omega,\mathscr{A})$ 上的有限测度, 并且每一 $\bar{\nu}_n$ 皆是 $\bar{\mu}_n$ 连续的.

i) $\frac{d\mu_1}{d\bar{\mu}_n}\to\frac{d\mu_1}{d\bar{\mu}},\bar{\mu}-\text{a.e.},$

ii) 如果 $\{\bar{\mu}_n\}$ 是 $\bar{\nu}$ 连续的, 则 $\dfrac{d\bar{\mu}_n}{d\bar{\nu}} \to \dfrac{d\bar{\mu}}{d\bar{\nu}}, \bar{\nu}-\text{a.e.}$,

iii) $\bar{\nu}$ 是 $\bar{\mu}$ 连续的, 并且 $\dfrac{d\bar{\nu}_n}{d\bar{\mu}_n} \to \dfrac{d\bar{\nu}}{d\bar{\mu}}, \bar{\mu}-\text{a.e.}$, (提示: 关于最后一个命题注意若对所有的 $n, \bar{\mu}_n(A_n) = 0$, 则 $\bar{\mu}(\overline{\lim\limits_{n\to\infty}} A_n) = 0$, 由此可知, 只要考察一个特殊选择的 $\dfrac{d\bar{\nu}_n}{d\bar{\mu}_n} = \sum\limits_{k=1}^{n} f_k \Big/ \sum\limits_{k=1}^{n} g_k$, 其中 $f_k = \dfrac{d\nu_k}{d\bar{\mu}}, g_k = \dfrac{d\mu_k}{d\bar{\mu}}$, 但是 $\sum f_n = \dfrac{d\bar{\nu}}{d\bar{\mu}}$, 并且 $\sum g_n = 1, \bar{\mu}-\text{a.e.}$)

7. 若 $(\Omega, \mathscr{A}, \mu)$ 是 σ-有限测度空间, 欲使 $\mathscr{A}$ 上的函数 φ 是 $(\Omega, \mathscr{A})$ 可测函数 f 关于 μ 的不定积分, 必须且只须 φ 是 σ-可加的并且对于每一集$A = \{a \leqslant f \leqslant b\} \cap B, B \in \mathscr{A}$, 恒有

$$a\mu(A) \leqslant \varphi(A) \leqslant b\mu(A).$$

8. 完成定理 7 的全部证明.

9. 设 f 是 $(\Omega, \mathscr{A}, \mu)$ 上的可测函数, f 关于 μ 的积分存在, 令

$$\varphi(A) = \int_A f d\mu, \quad A \in \mathscr{A}.$$

试证

$$\varphi^+(A) = \int_A f^+ d\mu, \quad \varphi^-(A) = \int_A f^- d\mu,$$

$$\bar{\varphi}(A) = \int_A |f| d\mu.$$

10. 设 φ 是 $(\Omega, \mathscr{A})$ 上 σ-可加集函数, f 是使得下式右端有意义的可测函数, 定义

$$\int f d\varphi = \int f d\varphi^+ - \int f d\varphi^-.$$

试证这个积分具有 § 2 所述的主要性质.

11. 集代数 $\mathscr{F}$ 上的 σ-可加集函数 φ 可以扩张为 $\sigma(\mathscr{F})$ 上的 σ-可加集函数的充分必要条件是 φ 是有下界或有上界的. 并且, 如果 φ 是 σ-有限的, 则扩张是唯一的, σ-有限的.

12. 若 μ 不是 σ-有限的, 则即使 φ 有限, Radon-Nikodym 定理也不一定成立. (提示: 考虑反例: $\Omega = [0,1], \mathscr{A} = \{A : A \subset [0,1],\ A\text{或}A^c\text{可列}\}$, $\mu(A) = A$ 中点的个数, $\varphi(A) = \begin{cases} 0, A\text{可列}, \\ 1, A^c\text{可列}. \end{cases}$)

第 4 章　乘积测度空间

在第 1 章 §1.3 我们已经给出了乘积空间与乘积 σ-代数的定义，并且在构成 n 维 Borel 域和 n 维实 (或复) 可测函数的过程中见到了它们的应用. 这一章我们要在乘积空间上建立一种测度 —— 乘积测度, 并且研究在乘积空间上可测函数关于乘积测度的积分.

首先我们从概率论和数理统计的实际需要来阐明建立乘积测度空间的必要性.

在数理统计中, 一个基本思想是借助于随机变量 ξ 在 n 次独立试验中的观测值 $x_1,\cdots,x_n$ 来推测 ξ 的概率性质. 由于在每次试验中所得到的观测值具有随机性, 因此我们将这些观测值看作是同 ξ 具有相同分布律的 n 个随机变量, 记为 $\xi_1,\cdots,\xi_n$; 又因为试验是独立进行的, 所以 $\xi_1,\cdots,\xi_n$ 应该是独立的. 于是它们的分布律应为

$$P(\xi_1\in B_1,\cdots,\xi_n\in B_n)=P(\xi_1\in B_1)\cdots P(\xi_n\in B_n). \tag{1}$$

上面的叙述看来是自然而合理的, 但仔细分析却有待于严格化, 其原因在于 $\xi_1,\cdots,\xi_n$ 是我们根据直观要求而引入的随机变量. 我们必须证明存在一个概率场, 在这个概念场上存在 n 维随机变量 $(\xi_1,\cdots,\xi_n)$, 其分布律满足 (1) 的要求, 因而 $\xi_1,\cdots,\xi_n$ 是独立的.

现在我们来考虑解决上述问题的途径.

设 ξ 是概率场 $(\Omega,\mathscr{A},P)$ 上的随机变量. 在直观上;$(\xi_1,\cdots,\xi_n)$ 的值是 ξ 在 n 次独立试验中依次出现的值 $(x_1,\cdots,x_n)$, 而 $(x_1,\cdots,x_n)$ 又反映了在 n 次独立试验中依次出现了基本事件 $\omega_1,\cdots,\omega_n$, 其中 $\xi(\omega_1)=x_1,\cdots,\xi(\omega_n)=x_n$, 因而自然我们取

$$\Omega^{(n)}=\Omega\times\cdots\times\Omega=\{(\omega_1,\cdots,\omega_n):\omega_k\in\Omega,k=1,\cdots,n\},$$

它是 $\Omega,\cdots,\Omega$(共 n 个) 的乘积空间. 以它为基本事件集合, 而定义

$$\xi_k(\omega_1,\cdots,\omega_n)=\xi(\omega_k),\qquad k=1,\cdots,n.$$

这样, $(\xi_1,\cdots,\xi_n)$ 可以满足我们提出的用来反映 ξ 在 n 次独立试验中所取的值的要求, 此外我们还应在 $\Omega^{(n)}$ 上构造一个 σ-代数 $\mathscr{A}^{(n)}$, 使 $\xi_1,\cdots,\xi_n$ 是关于它可测的. 由于

$$\{\xi_1\in B_1,\cdots,\xi_n\in B_n\}$$

$$
\begin{aligned}
&=\{(\omega_1,\cdots,\omega_n):\xi_i(\omega_1,\cdots,\omega_n)\in B_i,i=1,\cdots,n\}\\
&=\{(\omega_1,\cdots,\omega_n):\omega_i\in A_i,i=1,\cdots,n\}\\
&=A_1\times\cdots\times A_n,
\end{aligned}
$$

其中

$$A_i=\{\omega_i:\omega_i\in\Omega,\xi(\omega_i)=\xi_i(\omega_1,\cdots,\omega_n)\in B_i\}=\{\xi\in B_i\}.$$

为使 $(\xi_1,\cdots,\xi_n)$ 关于 $\mathscr{A}^{(n)}$ 可测, 必须使 $\mathscr{A}^{(n)}$ 包含一切可测矩形 $A_1\times\cdots\times A_n$. 而这样的可测矩形类上的最小 σ-代数就是 $\mathscr{A}\times\cdots\times\mathscr{A}$(即 n 个 $\mathscr{A}$ 的乘积 σ-代数). 因而取

$$\mathscr{A}^{(n)}=\underbrace{\mathscr{A}\times\cdots\times\mathscr{A}}_{n\text{个}}.$$

在 $\mathscr{A}^{(n)}$ 上定义一个概率 $P^{(n)}$, 使它满足

$$P^{(n)}(A_1\times\cdots\times A_n)=P(A_1)\times\cdots\times P(A_n).$$

这样一来, 我们就把问题归结为下述一般问题：给了 n 个概率场 $(\Omega_k,\mathscr{A}_k,P_k),k=1,\cdots,n$, 是否可以建立一个概率场 $(\Omega^{(n)},\mathscr{A}^{(n)},P^{(n)})$, 使得

$$\Omega^{(n)}=\Omega_1\times\cdots\times\Omega_n,$$

$$\mathscr{A}^{(n)}=\mathscr{A}_1\times\cdots\times\mathscr{A}_n.$$

而 $P^{(n)}$ 满足条件

$$P^{(n)}(A_1\times\cdots\times A_n)=P_1(A_1)\cdots P_n(A_n),\ \text{对一切}\ A_k\in\mathscr{A}_k,$$

$$k=1,\cdots,n.$$

前面提到的问题就是这个一般问题中 $(\Omega_k,\mathscr{A}_k,P_k),k=1,\cdots,n$ 都相同的情形.

更一般些, 如果最初所提的问题中, 试验次数不限定在一个固定的次数以内, 而是可以任意地增加, 类似地我们就需要引入一个独立随机变量序列：$\xi_1,\cdots,\xi_n,\cdots$, 它们分别与 ξ 具有相同的分布律. 这样就导致下面更一般的问题：是否可以建立如下的概率场 $(\Omega^{(\infty)},\mathscr{A}^{(\infty)},F^{(\infty)})$, 使

$$\Omega^{(\infty)}=\{(\omega_1,\cdots,\omega_n,\cdots):\omega_n\in\Omega_n,n=1,2,\cdots\},$$

$\mathscr{A}^{(\infty)}$ 是包含一切下列形式的集合的最小 σ-代数：

$$A_1\times\cdots\times A_n\times\Omega_{n+1}\times\cdots=\{(\omega_1,\cdots,\omega_n,\cdots):\omega_k\in A_k,$$

$$k=1,\cdots,n,\omega_i\in\Omega_i,i=n+1,n+2,\cdots\},$$

其中 $A_k\in\mathscr{A}_k,k=1,\cdots,n,n$ 是任意自然数, $P^{(\infty)}$ 满足下面条件: (称为无穷乘积概率)

$$P^{(\infty)}(A_1\times\cdots\times A_n\times\Omega_{n+1}\times\cdots)=P_1(A_1)\times\cdots\times P_n(A_n),$$

其中 $A_k\in\mathscr{A}_k,k=1,\cdots,n,n$ 是任意自然数.

这一章里我们来讨论上面的问题, 与此密切相关的是: 既然 $(\Omega^{(n)},\mathscr{A}^{(n)})$ 上的概率由 $P_k,k=1,\cdots,n$ 来决定, 因此 $(\Omega^{(n)},\mathscr{A}^{(n)},P^{(n)})$ 上可测函数的积分也应该由 $(\Omega_k,\mathscr{A}_k,P_k),k=1,\cdots,n$ 上的积分表示, 这就是重积分与累次积分的互化问题 (Fubini 定理), 也将在本章讨论.

上面提出的一些问题, 有些可以在具体的测度空间上来讨论, 例如一维 Borel 集与 n 维 Borel 集, 直线上的 Lebesgue 测度与 n 维空间中的 Lebesgue 测度 (实际上是它在 Borel 域上的限制) 的关系, 就是上述一般问题的特例. 因而本章所讨论问题的重要性不限于概率论方面. 有限维乘积空间与 Fubini 定理有着广泛的应用. 但是本章仍然着眼于这些理论在概率论中的应用, 举出一些实例, 说明有限维乘积测度、Fubini 定理及无穷乘积概率是如何用来从理论上讲清一些实践中遇到的问题的.

§4.1　有限维乘积测度

我们首先建立二维乘积测度, 为此先给出以下概念.

定义 1　设$A\subset\Omega_1\times\Omega_2$, 对于任意取定的 $\omega_1\in\Omega_1$, 集合

$$A_{\omega_1}=A(\omega_1)=\{\omega_2:(\omega_1,\omega_2)\in A\}$$

称为**A在 ω_1 处的截集**; 对于任一取定的 $\omega_2\in\Omega_2$,

$$A_{\omega_2}=A(\omega_2)=\{\omega_1:(\omega_1,\omega_2)\in A\}$$

称为**A在 ω_2 处的截集**(其中 A_{ω_i} 或 $A(\omega_i),i=1,2$ 是我们规定的用来表示截集的记号).

显然, 当 $\omega_1\in\Omega_1$ 时,$A(\omega_1)\subset\Omega_2$, 而当 $\omega_2\in\Omega_2$ 时 $A(\omega_2)\subset\Omega_1$, 并且若 $A=\varnothing$, 则 $A(\omega_i)=\varnothing,i=1,2$.

截集具有下述性质.

性质 1　对于 $i=1,2$ 有

$$A\cap B=\varnothing\Rightarrow A_{\omega_i}\cap B_{\omega_i}=\varnothing;\tag{2}$$

$$A \supset B \Longrightarrow A_{\omega_i} \supset B_{\omega_i}; \tag{3}$$

$$A = \bigcup_m A^{(m)} \Longrightarrow A_{\omega_i} = \bigcup_m A_{\omega_i}^{(m)}; \tag{4}$$

$$A = \bigcap_m A^{(m)} \Longrightarrow A_{\omega_i} = \bigcap_m A_{\omega_i}^{(m)}; \tag{5}$$

$$C = A - B \Longrightarrow C_{\omega_i} = A_{\omega_i} - B_{\omega_i}. \tag{6}$$

特别,

$$(A^c)_{\omega_i} = (A_{\omega_i})^c. \tag{7}$$

这些等式，读者根据定义不难验证.

定理 1 设$(\Omega_i, \mathscr{A}_i), i = 1, 2$ 是可测空间，若 $A \in \mathscr{A}_1 \times \mathscr{A}_2$, 则对任何 $\omega_1 \in \Omega_1$ 来说，$A_{\omega_1} \in \mathscr{A}_2$ 而对任何 $\omega_2 \in \Omega_2$ 来说，$A_{\omega_2} \in \mathscr{A}_1$.

这个定理也可以简单地叙述为：可测集 ($\mathscr{A}_1 \times \mathscr{A}_2$-可测) 的任何截集是可测集（相应地，$\mathscr{A}_1$-可测或 $\mathscr{A}_2$-可测).

证 设

$$\mathfrak{M} = \{A \in \mathscr{A}_1 \times \mathscr{A}_2 : \text{对一切}\omega_1 \in \Omega_1, A_{\omega_1} \in \mathscr{A}_2, \text{对一切}\omega_2 \in \Omega_2, A_{\omega_2} \in \mathscr{A}_1\},$$

$$\mathscr{C} = \{A = A_1 \times A_2 : A_1 \in \mathscr{A}_1, A_2 \in \mathscr{A}_2\}.$$

根据 $\mathscr{A}_1 \times \mathscr{A}_2$ 的定义, $\mathscr{A}_1 \times \mathscr{A}_2 = \sigma(\mathscr{C})$, 今往证 $\mathfrak{M}$ 是包含 $\mathscr{C}$ 的 σ-代数.

若 $A = A_1 \times A_2 \in \mathscr{C}$, 显然有

$$(A_1 \times A_2)_{\omega_1} = \begin{cases} A_2, & \omega_1 \in A_1, \\ \varnothing, & \omega_1 \notin A_1, \end{cases}$$

$$(A_1 \times A_2)_{\omega_2} = \begin{cases} A_1, & \omega_2 \in A_2, \\ \varnothing, & \omega_2 \notin A_2, \end{cases}$$

故 $\mathfrak{M} \supset \mathscr{C}$.

由性质 1 易证 $\mathfrak{M}$ 为 σ-代数, 故 $\mathfrak{M} \supset \sigma(\mathscr{C}) = \mathscr{A}_1 \times \mathscr{A}_2$, 而 $\mathfrak{M} \subset \mathscr{A}_1 \times \mathscr{A}_2$, 故 $\mathfrak{M} = \mathscr{A}_1 \times \mathscr{A}_2$. □

定义 2 设 $f(\omega_1, \omega_2)$ 是定义在 $\Omega_1 \times \Omega_2$ 上的函数. 对于任意取定的 $\omega_1 \in \Omega_1$, 函数

$$f_{\omega_1}(\omega_2) = f(\omega_1, \omega_2), \omega_2 \in \Omega_2,$$

作为 Ω_2 上的函数, 称为 **f 在 ω_1 处的截函数**; 对于任意取定的 $\omega_2 \in \Omega_2$, 函数

$$f_{\omega_2}(\omega_1) = f(\omega_1, \omega_2), \omega_1 \in \Omega_1,$$

作为 Ω_1 上的函数, 称为 f 在 ω_2 处的截函数.

定理 2　如果 f 是 $(\Omega_1\times\Omega_2,\mathscr{A}_1\times\mathscr{A}_2)$ 上的可测函数 (实的或复的), 则对任意取定的 $\omega_1\in\Omega_1, f_{\omega_1}$ 是 $(\Omega_2,\mathscr{A}_2)$ 上的可测函数, 而对任意取定的 $\omega_2\in\Omega_2, f_{\omega_2}$ 是 $(\Omega_1,\mathscr{A}_1)$ 上的可测函数. 简言之, 即可测函数的截函数是可测函数.

证　对于任意的 $B\in\mathscr{B}$(当 f 是实可测函数时 $\mathscr{B}=\tilde{\mathscr{B}}^{(1)}$, 当 f 是复可测函数时 $\mathscr{B}=\tilde{\mathscr{B}}_Z^{(1)}$),

$$f^{-1}(B)=\{(\omega_1,\omega_2):f(\omega_1,\omega_2)\in B\}\in\mathscr{A}_1\times\mathscr{A}_2,$$

而

$$\begin{aligned}f_{\omega_1}^{-1}(B)&=\Big\{\omega_2:f_{\omega_1}(\omega_2)\in B\Big\}=\Big\{\omega_2:f(\omega_1,\omega_2)\in B\Big\}\\&=\Big\{\omega_2:(\omega_1,\omega_2)\in f^{-1}(B)\Big\}=[f^{-1}(B)]_{\omega_1}.\end{aligned}$$

由定理 1 知 $f_{\omega_1}^{-1}(B)\in\mathscr{A}_2$, 故 f_{ω_1} 是 $(\Omega_2,\mathscr{A}_2)$ 上的可测函数. 同理可证 f_{ω_2} 是 $(\Omega_1,\mathscr{A}_1)$ 上的可测函数. □

定理 3　设 f 是 $(\Omega_1\times\Omega_2,\mathscr{A}_1\times\mathscr{A}_2)$ 上的非负可测函数,μ_i 是 $(\Omega_i,\mathscr{A}_i), i=1,2$ 上的 σ-有限测度, 则

$$\int_{\Omega_1}f(\omega_1,\omega_2)d\mu_1(\omega_1)\left(\text{相应地}\int_{\Omega_2}f(\omega_1,\omega_2)d\mu_2(\omega_2)\right)$$

是 $(\Omega_2,\mathscr{A}_2)$(相应地 $(\Omega_1,\mathscr{A}_1)$) 上的可测函数.

特别, 对一切 $A\in\mathscr{A}_1\times\mathscr{A}_2, \mu_1(A(\omega_2))$(相应地 $\mu_2(A(\omega_1))$) 是 $\mathscr{A}_2$(相应地 $\mathscr{A}_1$) 可测函数.

证　由于证法完全相同, 我们只证关于 $\int_{\Omega_1}f(\omega_1,\omega_2)d\mu_1(\omega_1)$ 的论断.

若定理的前一部分获证, 则令 $f(\omega_1,\omega_2)=\chi_A(\omega_1,\omega_2)$ 即得后一部分结论. 再者, 由于对一切 $\omega_2\in\Omega_2$ 来说,$f_{\omega_2}(\omega_1)$ 是 $\mathscr{A}_1$ 可测的, 故 $\int_{\Omega_1}f(\omega_1,\omega_2)d\mu_1(\omega_1)$ 存在.

i) 先证对任何 $A\in\mathscr{A}_1\times\mathscr{A}_2, \int_{\Omega_1}\chi_A(\omega_1,\omega_2)d\mu_1(\omega_1)$ 是 $\mathscr{A}_2$ 可测的.

由于 μ_1 是 $(\Omega_1,\mathscr{A}_1)$ 上 σ-有限测度, 因而存在 $\Omega_1^{(n)}\in\mathscr{A}_1, n=1,2,\cdots$ 两两不交,$\mu_i(\Omega_1^{(n)})<\infty, n=1,2,\cdots$ 且 $\sum\limits_{n=1}^{\infty}\Omega_1^{(n)}=\Omega_1$. 令

$$\mathfrak{M}=\left\{A:\text{对一切正整数}n,\int_{\Omega_1^{(n)}}\chi_A(\omega_1,\omega_2)d\mu_1(\omega_1)\text{是}\mathscr{A}_2\text{-可测的}\right\}.$$

易知, 若 $A\in\mathfrak{M}$, 则 $\int_{\Omega_1}\chi_A(\omega_1,\omega_2)d\mu_1(\omega_1)=\sum_{n=1}^{\infty}\int_{\Omega_1^{(n)}}\chi_A(\omega_1,\omega_2)\cdot d\mu_1(\omega_1)$ 是 $\mathscr{A}_2$-可测的.

令

$$\mathscr{C}=\{A_1\times A_2:A_i\in\mathscr{A}_i,i=1,2\}.$$

$\mathscr{C}$ 是 π 系, 若 $A\in\mathscr{C},A=A_1\times A_2,A_i\in\mathscr{A}_i,i=1,2$, 对任何正整数 n, $\int_{\Omega_1^{(n)}}\chi_A(\omega_1,\omega_2)d\mu_1(\omega_1)=\chi_{A_2}(\omega_2)\mu_1(A_1\cap\Omega_1^{(n)})$ 是 $\mathscr{A}_2$-可测的, 故 $\mathscr{C}\subset\mathfrak{M}$.

再证 $\mathfrak{M}$ 是 λ-系. 因为 $\Omega_1\times\Omega_2\in\mathscr{C}$, 因而 $\Omega_1\times\Omega_2\in\mathfrak{M}$. 若 $A,B\in\mathfrak{M},A\subset B$, 即对一切正整数 n, $\int_{\Omega_1^{(n)}}\chi_A(\omega_1,\omega_2)d\mu_1(\omega_1)$, $\int_{\Omega_1^{(n)}}\chi_B(\omega_1,\omega_2)d\mu_1(\omega_1)$ 是 $\mathscr{A}_2$ 可测的, 且易证它们都是有限的, 故对一切正整数 n,

$$\begin{aligned}\int_{\Omega_1^{(n)}}\chi_{B-A}(\omega_1,\omega_2)d\mu_1(\omega_1)&=\int_{\Omega_1^{(n)}}(\chi_B(\omega_1,\omega_2)-\chi_A(\omega_1,\omega_2))d\mu_1(\omega_1)\\&=\int_{\Omega_1^{(n)}}\chi_B(\omega_1,\omega_2)d\mu_1(\omega_1)-\int_{\Omega_1^{(n)}}\chi_A(\omega_1,\omega_2)d\mu_1(\omega_1)\end{aligned}$$

是 $\mathscr{A}_2$-可测函数。即 $\mathfrak{M}$ 对真差封闭。再若 $A_m\in\mathfrak{M},m=1,2,\cdots,A_m\uparrow A$, 则对任何正整数 n,

$$\int_{\Omega_1^{(n)}}\chi_A(\omega_1,\omega_2)d\mu_1(\omega_1)=\lim_{m\to\infty}\int_{\Omega_1^{(n)}}\chi_{A_m}(\omega_1,\omega_2)d\mu_1(\omega_1)$$

是 $\mathscr{A}_2$-可测的, 因而 $\mathfrak{M}$ 对单调增运算封闭. 因而 $\mathfrak{M}$ 为包含 π-系 $\mathscr{C}$ 的 λ-系, 故 $\mathfrak{M}\supset\sigma(\mathscr{C})=\mathscr{A}_1\times\mathscr{A}_2$. 这就证明了对任何 $A\in\mathscr{A}_1\times\mathscr{A}_2$, $\int_{\Omega_1}\chi_A(\omega_1,\omega_2)d\mu_1(\omega_1)$ 是 $\mathscr{A}_2$ 可测函数.

ii) 再证 f 是 $(\Omega_1\times\Omega_2,\mathscr{A}_1\times\mathscr{A}_2)$ 上任意非负可测函数的情形.

事实上, 由 i) 知当 $f(\omega_1,\omega_2)=\chi_A(\omega_1,\omega_2),A\in\mathscr{A}_1\times\mathscr{A}_2$ 时 $\int_{\Omega_1}f(\omega_1,\omega_2)d\mu_1(\omega_1)$ 是 $\mathscr{A}_2$-可测的. 于是由积分线性性质知当 f 是非负简单函数时定理的结论正确, 再由单调收敛定理知 f 是非负可测函数时定理成立. □

现在证明乘积测度的存在及唯一性定理.

定理 4 设 $(\Omega_i,\mathscr{A}_i,\mu_i),i=1,2$ 是两个 σ-有限测度空间, 若令

$$\mu(A)\triangleq\int_{\Omega_1}\mu_2(A(\omega_1))d\mu_1(\omega_1),A\in\mathscr{A}_1\times\mathscr{A}_2\tag{8}$$

或

$$\mu(A) \triangleq \int_{\Omega_2} \mu_1(A(\omega_2))d\mu_2(\omega_2), A \in \mathscr{A}_1 \times \mathscr{A}_2, \tag{9}$$

则 μ 是 $\mathscr{A}_1 \times \mathscr{A}_2$ 上唯一的一个满足

$$\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2), \text{对一切} A_i \in \mathscr{A}_i, i = 1, 2 \tag{10}$$

的 σ-有限测度.

证　i) 先证满足条件 (10) 的 μ 一定 σ-有限且唯一. 事实上, 由于 μ_i 是 σ-有限的, 故有 $\Omega_i^{(n)} \in \mathscr{A}_i, n = 1, 2, \cdots$ 两两不交, $\mu_i(\Omega_i^{(n)}) < \infty, n = 1, 2, \cdots$, 且 $\sum_{n=1}^{\infty} \Omega_i^{(n)} = \Omega_i, i = 1, 2$. 于是 $\Omega_1^{(m)} \times \Omega_2^{(n)} \in \mathscr{A}_1 \times \mathscr{A}_2, m, n = 1, 2, \cdots$ 两两不交, $\sum_{m,n} \Omega_1^{(m)} \times \Omega_2^{(n)} = \Omega_1 \times \Omega_2$, 且对每一 m, n

$$\mu(\Omega_1^{(m)} \times \Omega_2^{(n)}) = \mu_1(\Omega_1^{(m)}) \cdot \mu_2(\Omega_2^{(n)}) < \infty.$$

因而 μ 是 σ-有限的. 其次, 若还有 μ' 满足 (10), 即

$$\mu'(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2), A_i \in \mathscr{A}_i, i = 1, 2,$$

则

$$\mu'(A_1 \times A_2) = \mu(A_1 \times A_2), A_i \in \mathscr{A}_i, i = 1, 2,$$

即 μ' 与 μ 在可测矩形类 $\mathscr{C}$ 上相等, 但 $\mathscr{C}$ 是半集代数,$\mathscr{A}_1 \times \mathscr{A}_2 = \sigma(\mathscr{C})$, 再由测度扩张定理知 μ' 与 μ 在 $\mathscr{A}_1 \times \mathscr{A}_2$ 上相等.

ii) 令

$$\mu(A) = \int_{\Omega_1} \mu_2(A(\omega_1))d\mu_1, \quad A \in \mathscr{A}_1 \times \mathscr{A}_2,$$

由定理 3 知对一切 $A \in \mathscr{A}_1 \times \mathscr{A}_2, \mu_2(A(\omega_1))$ 是 $\mathscr{A}_1$-可测的, 显然它是非负的, 因而由 (8) 定义的 $\mu(A)$ 是 $\mathscr{A}_1 \times \mathscr{A}_2$ 上的非负集函数.

再证 μ 具有 σ-可加性.

设 $A^{(n)} \in \mathscr{A}_1 \times \mathscr{A}_2, n = 1, 2, \cdots$ 两两不交, 则由

$$\left(\sum_m A^{(n)}\right)(\omega_1) = \sum_n A^{(n)}(\omega_1),$$

且 $A^{(n)}(\omega_1) \in \mathscr{A}_2, n = 1, 2, \cdots$ 两两不交, 故由 μ_2 的 σ-可加性及单调收敛定理知

$$\begin{aligned}\mu\left(\sum_n A^{(n)}\right) &= \int_{\Omega_1} \mu_2\left(\sum_n A^{(n)}(\omega_1)\right) d\mu_1 \\ &= \int_{\Omega_1} \sum_n \mu_2(A^{(n)}(\omega_1)) d\mu_1 \\ &= \sum_n \int_{\Omega_1} \mu_2(A^{(n)}(\omega_1)) d\mu_1 \\ &= \sum_n \mu(A^{(n)}).\end{aligned}$$

故 μ 是 $\mathscr{A}_1 \times \mathscr{A}_2$ 上的一个测度.

由于对任何 $A_1 \times A_2 \in \mathscr{C}$,

$$\begin{aligned}\mu(A_1 \times A_2) &= \int_{\Omega_1} \mu_2((A_1 \times A_2)(\omega_1)) d\mu_1 \\ &= \int_{\Omega_1} \mu_2(A_2)\chi_{A_1}(\omega_1) d\mu_1 = \mu_1(A_1) \cdot \mu_2(A_2),\end{aligned}$$

故由 (8) 定义的 μ 满足 (10). 同样可证由 (9) 定义的 μ 也满足 (10). 由 i) 知 (8)(或 (9)) 定义的 μ 是满足 (10) 的 $\mathscr{A}_1 \times \mathscr{A}_2$ 上唯一的测度, 且它是 σ-有限的. □

定义 3 设 $(\Omega_i, \mathscr{A}_i, \mu_i), i = 1, 2$ 是两个 σ-有限测度空间, 称满足 (10) 的 μ 为 μ_1 和 μ_2 的**乘积测度**, 记作 $\mu_1 \times \mu_2$, 而 $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \times \mathscr{A}_2, \mu_1 \times \mu_2)$ 称为 $(\Omega_i, \mathscr{A}_i, \mu_i), i = 1, 2$ 的乘积测度空间.

推论 设 $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \times \mathscr{A}_2, \mu_1 \times \mu_2)$ 是 $(\Omega_i, \mathscr{A}_i, \mu_i), i = 1, 2$ 的乘积测度空间, 则 $A \in \mathscr{A}_1 \times \mathscr{A}_2, \mu(A) = 0$ 的充分必要条件是

$$\mu_2(A(\omega_1)) = 0, \quad \text{a.e.}\, \mu_1$$

或

$$\mu_1(A(\omega_2)) = 0, \quad \text{a.e.}\, \mu_2.$$

证 由于 $\mu_2(A(\omega_1))$ 是非负 $\mathscr{A}_1$-可测的, 由积分性质知

$$\mu(A) = \int_{\Omega_1} \mu_2(A(\omega_1)) d\mu_1 = 0 \Longleftrightarrow \mu_2(A(\omega_1)) = 0 (\text{a.e.}\mu_1).$$

同理可证

$$\mu(A) = \int_{\Omega_2} \mu_1(A(\omega_2)) d\mu_2 = 0 \Longleftrightarrow \mu_1(A(\omega_2)) = 0 (\text{a.e.}\, \mu_2).$$ □

例 1 设 $(R^{(1)}, \mathscr{B}^{(1)}, \mu_1), i = 1, 2$ 是两个一维 L-S 测度空间, 设 μ_i 是由分布函数 $F_i(x)$ 决定的, $i = 1, 2$. 则由上述定理知有 $(R^{(1)} \times R^{(1)}, \mathscr{B}^{(1)} \times \mathscr{B}^{(1)}, \mu_1 \times \mu_2)$(即

$(R^{(2)}, \mathscr{B}^{(2)}, \mu_1 \times \mu_2))$ 存在, 且 $\mu = \mu_1 \times \mu_2$ 是 σ-有限测度. 今决定 μ 的分布函数. 由于对一切 $a = (a_1, a_2), b = (b_1, b_2)$,

$$[a, b) = [a_1, b_1) \times [a_2, b_2).$$

故

$$\begin{aligned}\mu([a, b)) &= \mu_1([a_1, b_1)) \cdot \mu_2([a_2, b_2)) \\ &= (F_1(b_1) - F_1(a_1)) \cdot (F_2(b_2) - F_2(a_2)) \\ &= F_1(b_1)F_2(b_2) - F_1(b_1)F_2(a_2) \\ &\quad -F_1(a_1)F_2(b_2) + F_1(a_1)F_2(a_2),\end{aligned}$$

若令

$$F(x_1, x_2) = F_1(x_1)F_2(x_2),$$

则

$$\mu([a, b)) = \Delta_{b,a}F.$$

因而 μ 是 $\mathscr{B}^{(2)}$ 上的 L-S 测度, 且它由分布函数 $F(x_1, x_2) = F_1(x_1)F_2(x_2)$ 唯一决定.

对于有限维乘积可测空间上乘积测度的构造, 可以使用归纳法 (也可以完全比照二维的情形) 来证明一系列的定理.

定义 1′　设 $A \subset \Omega_1 \times \cdots \times \Omega_n$, 则对任何 $(\omega_{i_1}, \cdots, \omega_{i_k}) \in \Omega_{i_1} \times \cdots \times \Omega_{i_k}$,

$$A(\omega_{i_1}, \cdots, \omega_{i_k}) = \{(\omega_{j_1}, \cdots, \omega_{j_{n-k}}) : (\omega_1, \cdots, \omega_n) \in A\}$$

称为 A 在 $(\omega_{i_1}, \cdots, \omega_{i_k})$ 处的截集. (其中 $j_1 < j_2 < \cdots < j_{n-k}, i_1 < \cdots < i_k$ 且 $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$ 是 $(1, \cdots, n)$ 的一个置换.)

定义 2′　设 $f(\omega_1, \cdots, \omega_n)$ 是定义在 $\Omega_1 \times \cdots \times \Omega_n$ 上的函数, 对于任意取定的 $(\omega_{i_1}, \cdots, \omega_{i_k}) \in \Omega_{i_1} \times \cdots \times \Omega_{i_k}$, 函数

$$f_{(\omega_{i_1}, \cdots, \omega_{i_k})}(\omega_{j_1}, \cdots, \omega_{j_{n-k}}) = f(\omega_1, \cdots, \omega_n),$$

$$(\omega_{j_1}, \cdots, \omega_{j_{n-k}}) \in \Omega_{j_1} \times \cdots \times \Omega_{j_{n-k}}$$

称为 f 在 $(\omega_{i_1}, \cdots, \omega_{i_k})$ 处的截函数. 其中 $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$ 是 $(1, \cdots, n)$ 的一个置换, 且 $i_1 < \cdots < i_k, j_1 < \cdots < j_{n-k}$.

定理 1 和定理 2 仍然成立. 即可测集的截集是可测集, 即若 $A \in \mathscr{A}_1 \times \cdots \times \mathscr{A}_n$, 则对 $(1, \cdots, n)$ 的任一置换 $(i_1, \cdots, i_k, j_1, \cdots j_{n-k})$(其中 $i_1 < \cdots < i_k, j_1 < \cdots <$

j_{n-k}), 及 $\omega_{i_1} \in \Omega_{i_1}, \cdots, \omega_{i_k} \in \Omega_{i_k}, A(\omega_{i_1}, \cdots, \omega_{i_k}) \in \mathscr{A}_{j_1} \times \cdots \times \mathscr{A}_{j_{n-k}}$; 可测函数的截函数仍然是可测函数. 即若 f 是 $\mathscr{A}_1 \times \cdots \times \mathscr{A}_n$ 可测函数, 则对 $(1, \cdots, n)$ 的任一置换 $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$ (其中 $i_1 < \cdots < i_k, j_1 < \cdots < j_{n-k}$) 及 $\omega_{i_1} \in \Omega_{i_1}, \cdots, \omega_{i_k} \in \Omega_{i_k}, f_{(\omega_{i_1}, \cdots, \omega_{i_k})}$ 作为 $\Omega_{j_1} \times \cdots \times \Omega_{j_{n-k}}$ 上的函数是 $\mathscr{A}_{j_1} \times \cdots \times \mathscr{A}_{j_{n-k}}$ 可测的.

定理 3 的类比是

定理 3′ 设 f 是 $(\Omega_1 \times \cdots \times \Omega_n, \mathscr{A}_1 \times \cdots \times \mathscr{A}_n)$ 上的非负可测函数, μ_i 是 $(\Omega_i, \mathscr{A}_i), i = 1, \cdots, n$ 上的 σ-有限测度, 则对任何 $(i_1, \cdots, i_k) \subset (1, \cdots, n)$,

$$\int_{\Omega_{i_k}} \cdots \int_{\Omega_{i_1}} f(\omega_1, \cdots, \omega_n) d\mu_{i_1}(\omega_{i_1}) \cdots d\mu_{i_k}(\omega_{i_k})$$

是 $(\Omega_{j_1} \times \cdots \times \Omega_{j_{n-k}}, \mathscr{A}_{j_1} \times \cdots \times \mathscr{A}_{j_{n-k}})$ 上的可测函数. 其中 $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$ 是 $(1, \cdots, n)$ 的一个置换且 $j_1 < \cdots < j_{n-k}$.

在 n 维乘积可测空间上的乘积测度存在唯一性定理仍然成立, 即

定理 4′ 设 $(\Omega_i, \mathscr{A}_i, \mu_i)$ 是 n 个 σ-有限测度空间, 则在乘积空间 $(\Omega_1 \times \cdots \times \Omega_n, \mathscr{A}_1 \times \cdots \times \mathscr{A}_n)$ 上存在唯一的满足条件

$$\mu(A_1 \times \cdots \times A_n) = \mu_1(A_1) \cdots \mu_n(A_n), A_i \in \mathscr{A}_i, \quad i = 1, \cdots, n \tag{11}$$

的测度 μ, 称此测度为乘积测度, 记作 $\mu_1 \times \cdots \times \mu_n$ 或 $\prod_{i=1}^{n} \mu_i$, 且 μ 是 σ-有限的. 它可以用累次积分得到, 即

$$\mu_1 \times \cdots \times \mu_n(A) = \int_{\Omega_{i_n}} \cdots \left(\int_{\Omega_{i_1}} \chi_A(\omega_1, \cdots, \omega_n) d\mu_i, (\omega_{i_1}) \right) \cdots d\mu_{i_n}(\omega_{i_n}), \tag{12}$$

其中 $(i_1, \cdots, i_n)$ 是 $(1, \cdots, n)$ 的任一置换.

证 我们先利用数学归纳法证明存在一个 $\mathscr{A}_1 \times \cdots \times \mathscr{A}_n$ 上的 σ-有限测度 μ 满足 (11). 事实上, 当 $n = 2$ 时定理成立 (定理 4 所证), 假定定理对于 $n - 1$ 成立, 则存在一个 $\mathscr{A}_1 \times \cdots \times \mathscr{A}_{n-1}$ 上的 σ-有限测度 μ' 满足

$$\mu'(A_1 \times \cdots \times A_{n-1}) = \mu_1(A_1) \cdots \mu_{n-1}(A_{n-1}), \quad A_i \in \mathscr{A}_i, i = 1, \cdots, n-1. \tag{13}$$

由此利用定理 4 推知, 存在一个 $\mathscr{A}_1 \times \cdots \times \mathscr{A}_n = (\mathscr{A}_1 \times \cdots \times \mathscr{A}_{n-1}) \times \mathscr{A}_n$ 上的测度 μ 满足

$$\mu(A' \times A_n) = \mu'(A')\mu_n(A_n), A' \in \mathscr{A}_1 \times \cdots \times \mathscr{A}_{n-1},$$

$$A_n \in \mathscr{A}_n. \tag{14}$$

由 (13),(14), 这个 μ 满足

$$\mu(A_1 \times \cdots \times A_n) = \mu'(A_1 \times \cdots \times A_{n-1}) \cdot \mu_n(A_n)$$

$$= \mu_1(A_1) \cdots \mu_{n-1}(A_{n-1}) \mu_n(A_n), A_i \in \mathscr{A}_i, i = 1, \cdots, n.$$

这就证明了存在一个 $\mathscr{A}_1 \times \cdots \times \mathscr{A}_n$ 上的测度满足 (11).

关于满足 (11) 的测度的唯一性, 可以仿照定理 4 的方法证明.

取 $A = A_1 \times \cdots \times A_n$, 易知 (12) 右边定义的集函数满足 (11) 且具有非负性、σ-可加性, 因而是满足 (11) 的测度, 再由 $\mu_1 \times \cdots \times \mu_n$ 的唯一性知 (12) 成立.

定义 3′　设 $(\Omega_i, \mathscr{A}_i, \mu_i), i = 1, \cdots, n$ 是 n 个 σ-有限测度空间, 称 $(\Omega_1 \times \cdots \times \Omega_n, \mathscr{A}_1, \times \cdots \times \mathscr{A}_n, \mu_1 \times \cdots \times \mu_n)$ 为 $(\Omega_i, \mathscr{A}_i, \mu_i), i = 1, \cdots, n$ 的乘积测度空间, 其中 $\mu_1 \times \cdots \times \mu_n$ 为满足 (11) 的唯一的测度.

利用乘积测度的唯一性以及 $\mathscr{A}_1 \times \cdots \times \mathscr{A}_n = (\mathscr{A}_1 \times \cdots \times \mathscr{A}_k) \times (\mathscr{A}_{k+1} \times \cdots \times \mathscr{A}_n), k = 2, \cdots, n-1$, 可以得到

推论　对任何 $k = 2, \cdots, n-1$,

$$(\mu_1 \times \cdots \times \mu_k) \times (\mu_{k+1} \times \cdots \times \mu_n) = \mu_1 \times \cdots \times \mu_n, \tag{15}$$

由此推论, 利用定理 4 可得

$$(\mu_1 \times \cdots \times \mu_n)(A)$$

$$= \int_{\Omega_1 \times \cdots \times \Omega_k} (\mu_{k+1} \times \cdots \times \mu_n)(A(\omega_1, \cdots, \omega_k))$$

$$d\mu_1 \times \cdots \times \mu_k, \tag{16}$$

对一切 $A \in \mathscr{A}_1 \times \cdots \times \mathscr{A}_n$ 成立.

注意: 在这一节我们所定义的乘积测度, 有时称为独立乘积测度. 独立是概率论的术语, 而 (11) 类似于概率论中事件独立的定义. 在乘积可测空间 $(\Omega_1 \times \cdots \times \Omega_n, \mathscr{A}_1 \times \cdots \times \mathscr{A}_n)$ 上还可以构造非独立乘积测度, 关于这个问题, 在第 5 章中将会进行探讨.

现在我们举例说明有限维乘积测度在概率论中的应用.

我们先解决引言中提出的问题.

例 2　设 ξ 是概率场 $(\Omega, \mathscr{A}, P)$ 上的随机变量, 分布函数为 $F(x)$, 则存在一个概率场, 在其上有 n 个独立随机变量 $\xi_1, \cdots, \xi_n$, 它们分别与 ξ 具有相同的分布函数.

证 首先建立乘积概率空间, 令

$$\Omega^{(n)} = \underbrace{\Omega \times \cdots \times \Omega}_{n个}, \quad \mathscr{A}^{(n)} = \underbrace{\mathscr{A} \times \cdots \times \mathscr{A}}_{n个},$$

$$P^{(n)} = \underbrace{P \times \cdots \times P}_{n个}.$$

在 $\Omega^{(n)}$ 上定义

$$\xi_i(\omega_1, \cdots, \omega_n) = \xi(\omega_i), i = 1, \cdots, n.$$

由于

$$\begin{aligned}&\{\xi_1 < x_1, \cdots, \xi_n < x_n\}\\ =&\{(\omega_1, \cdots, \omega_n) : \xi(\omega_i) < x_i, i = 1, \cdots, n\}\\ =&\{\omega_1 : \xi(\omega_1) < x_1\} \times \cdots \times \{\omega_n : \xi(\omega_n) < x_n\}\end{aligned}$$

是可测矩形, 因而

$$P^{(n)}(\xi_1 < x_1, \cdots, \xi_n < x_n) = P(\xi < x_1) \cdots P(\xi < x_n),$$

即

$$F_{\xi_1, \cdots, \xi_n}(x_1, \cdots, x_n) = F(x_1) \cdots F(x_n).$$

由于 $(\xi_1, \cdots, \xi_n)$ 的分布函数可以分解成 n 个分布函数的乘积, 因此 $\xi_1, \cdots, \xi_n$ 独立 (第 2 章 §2.4 定理 2). 并且每一 ξ_i 都与 ξ 具有相同的分布函数 $F(x)$.

例 3 设 ξ_i 是概率场 $(\Omega_i, \mathscr{A}_i, P_i)$ 上的随机变量, 其分布函数为 $F_i(x), i = 1, \cdots, n$. 在乘积概率空间 $(\Omega^{(n)}, \mathscr{A}^{(n)}, P^{(n)})$ 上定义

$$\eta_i(\omega_1, \cdots, \omega_n) = \xi_i(\omega_i), (\omega_1, \cdots, \omega_n) \in \Omega^{(n)}, i = 1, \cdots, n.$$

则 $\eta_1, \cdots, \eta_n$ 是独立随机变量, 且 η_i 与 ξ_i 同分布, $i = 1, \cdots, n$.

证 由于对一切 $(x_1, \cdots, x_n) \in R^{(n)}$,

$$\begin{aligned}&\{\eta_1 < x_1, \cdots, \eta_n < x_n\}\\ =&\{(\omega_1, \cdots, \omega_n) : \xi_i(\omega_i) < x_i, i = 1, \cdots, n\}\\ =&\{\omega_1 : \xi_1(\omega_1) < x_1\} \times \cdots \times \{\omega_n : \xi_n(\omega_n) < x_n\}.\end{aligned}$$

故

$$F_{\eta_1, \cdots, \eta_n}(x_1, \cdots, x_n) = P^{(n)}(\eta_1 < x_1, \cdots, \eta_n < x_n)$$

$$= P_1(\xi_1 < x_1) \cdots P_n(\xi_n < x_n) = F_1(x_1) \cdots F_n(x_n).$$

因而 $\eta_1, \cdots, \eta_n$ 独立且 η_i 与 ξ_i 具有相同的分布函数 $F_i(x), i = 1, \cdots, n$.

例 4　对于任意给定的 n 个概率分布函数 $F_i(x_i), i = 1, \cdots, n$, 总存在 n 个独立的随机变量 $\xi_1, \cdots, \xi_n$, 使得 ξ_i 的分布函数恰好是 $F_i(x_i), i = 1, \cdots, n$.

证　由 $F_i(x_i)$ 可决定一个概率场 $(R^{(1)}, \mathscr{B}^{(1)}, P_i), i = 1, \cdots, n$, 其中 $(R^{(1)}, \mathscr{B}^{(1)})$ 是 Borel 空间. 由它们构成的乘积概率空间是 $(R^{(n)}, \mathscr{B}^{(n)}, P_1 \times \cdots \times P_n)$, 在其上定义

$$\xi_i(x_1, \cdots, x_n) = x_i, i = 1, \cdots, n.$$

则 ξ_i 以 $F_i(x_i)$ 为分布函数, 且 $\xi_1, \cdots, \xi_n$ 独立.

例 5　用乘积概率空间建立 n 次独立试验的概率场. 设第 i 次试验中事件 A_i 发生的概率为 p, 于是第 i 次试验的概率场为 $(\Omega_i, \mathscr{A}_i, P_i)$, 其中

$$\Omega_i = \{\omega_i, \bar{\omega}_i\}, \quad \begin{array}{l} \omega_i\text{表示事件}A_i\text{发生}, \\ \bar{\omega}_i\text{表示事件}A_i\text{不发生}; \end{array}$$

$$\mathscr{A}_i = \{\varnothing, \{\omega_i\}, \{\bar{\omega}_i\}, \Omega_i\};$$

$$P_i(\{\omega_i\}) = p, P_i(\{\bar{\omega}_i\}) = 1 - p.$$

在其上定义随机变量 ξ_i 为

$$\xi_i(\omega_i) = 1, \quad \xi_i(\bar{\omega}_i) = 0.$$

n 次独立试验的概率场为 $(\Omega_1 \times \cdots \times \Omega_n, \mathscr{A}_1 \times \cdots \times \mathscr{A}_n, P_1 \times \cdots \times P_n)$, 在其上定义随机变量

$$\eta_i(\omega'_1, \cdots, \omega'_n) = \xi_i(\omega'_i), \omega'_j = \omega_j\text{或}\bar{\omega}_j, j = 1, \cdots, n,$$
$$i = 1, \cdots, n.$$

则 $\eta_1, \cdots, \eta_n$ 是独立随机变量, 且

$$P_1 \times \cdots \times P_n(\eta_i = 1) = P_i(\xi_i = 1) = p,$$

$$P_1 \times \cdots \times P_n(\eta_i = 0) = P_i(\xi_i = 0) = 1 - p.$$

若令

$$\eta = \sum_{i=1}^{n} \eta_i,$$

则 η 是乘积概率空间上的随机变量, 它表示 n 次独立试验中事件 A_i 发生的次数, 今求其概率分布.

由于 η_i 只取 0, 1 为值, 因而 η 只取 $0,1,\cdots,n$ 为值, 且

$$\begin{aligned}&P_1\times\cdots\times P_n(\eta=k)\\=&\sum_{1\leqslant i_1<\cdots<i_k\leqslant n}P_1\times\cdots\times P_n(\eta_{i_s}=1,s=1,\cdots,k,\\&\eta_t=0,t\neq i_s,1\leqslant t\leqslant n)\\=&\binom{n}{k}p^k(1-p)^{n-k},k=0,1,\cdots,n.\end{aligned}$$

注意, 我们记 $\{\eta=k\}$ 为 $A_n^{(k)}$, 而

$$A_n^{(k)}=\{(\omega_1',\cdots,\omega_n'):\text{在}\omega_1',\cdots,\omega_n'\text{中恰有 }k\text{ 个}$$
$$\omega_i,n-k\text{个}\bar{\omega}_i\},k=0,1,\cdots,n.$$

故 $P_1\times\cdots\times P_n(A_n^{(k)})=\binom{n}{k}p^k(1-p)^{n-k}$. 这一结果在 §3 中将要用到.

习题及补充

1. 试问 $\bar{\mathscr{A}_1}\times\bar{\mathscr{A}_2}=\overline{\mathscr{A}_1\times\mathscr{A}_2}$ 吗? 其中 $\bar{\mathscr{A}_i}$ 表示 $\mathscr{A}_i$ 对 μ_i 的完备化,$i=1,2$;$\overline{\mathscr{A}_1\times\mathscr{A}_2}$ 表示 $\mathscr{A}_1\times\mathscr{A}_2$ 对测度 $\mu_1\times\mu_2$ 的完备化.

2. 若 Ω 是一不可数集,$\mathscr{A}$ 是包含 Ω 中一切单点集的最小 σ-代数, 则 $\Omega\times\Omega$ 的对角线 $A=\{(\omega,\omega):\omega\in\Omega\}$ 不属于 $\mathscr{A}\times\mathscr{A}$, 但对任意 $\omega_1\in\Omega,A_{\omega_1}\in\mathscr{A}$, 对任意的 $\omega_2\in\Omega,A_{\omega_2}\in\mathscr{A}$. 其中

$$A_{\omega_1}=\{\omega_2:(\omega_1,\omega_2)\in A\},$$
$$A_{\omega_2}=\{\omega_1:(\omega_1,\omega_2)\in A\}.$$

这说明定理 1 的逆命题不成立.

3. 设 ξ_1, ξ_2 是独立随机变量, $F_1(x),F_2(x)$ 是它们的分布函数. 试应用乘积概率定理证明 $\xi_1+\xi_2$ 的分布函数是

$$F(x)=\int_{-\infty}^{\infty}F_1(x-y)dF_2(y).$$

4. 设 $(\Omega,\mathscr{A},P)$ 是概率空间, $f(t,\omega)$ 是 $\mathscr{B}^{(1)}\times\mathscr{A}$ 可测的. 若对每个 $t\in R,f(t,\cdot)$, a.e.(P) 有限, 试证

$$\mu(\{t:f(t,\omega)=\infty\})=0,\text{a.e.}(P),$$

其中 μ 表示 Lebesgue 测度 (提示: 令 $A=\{(t,\omega);f(t,\omega)=\infty\}$, 考虑 $\mu\times P(A)$)

5. 设 f 是 $(\Omega_1\times\cdots\times\Omega_n,\mathscr{A}_1\times\cdots\times\mathscr{A}_n)$ 上的可测函数, 试证: 对于任意取定的 $(\omega_{i_1},\cdots,\omega_{i_k})\in\Omega_{i_1}\times\cdots\times\Omega_{i_k},(k<n)$, f 在 $(\omega_{i_1},\cdots,\omega_{i_k})$ 处的截函数 $f_{(\omega_{i_1},\cdots,\omega_{i_k})}$

是 $(\Omega_{j_1}\times\cdots\times\Omega_{j_{n-k}},\mathscr{A}_{j_1}\times\cdots\times\mathscr{A}_{j_{n-k}})$ 上的可测函数. 其中 $(i_1,\cdots,i_k,j_1,\cdots,j_{n-k})$ 是 $(1,\cdots,n)$ 的一个置换.

§4.2　Fubini 定理

这一节我们要讨论乘积测度空间 $(\Omega_1\times\cdots\times\Omega_n,\mathscr{A}_1\times\cdots\times\mathscr{A}_n,\mu_1\times\cdots\times\mu_n)$ 上的积分 (称为重积分) 与原来给定的测度空间 $(\Omega_i,\mathscr{A}_i,\mu_i),i=1,\cdots,n$ 上的积分的关系. 我们将证明, 在一定条件下重积分可以化成对 $\mu_1\cdots,\mu_n$ 的累次积分来进行计算, 通常称为 Fubini 定理, 它有着广泛的应用.

首先, 我们讨论 $n=2$ 的情形.

定理 1　设 $(\Omega_i,\mathscr{A}_i,\mu_i),i=1,2$ 是两个 σ-有限测度空间, 若 f 是 $\mathscr{A}_1\times\mathscr{A}_2$-可测的非负函数, 则

$$\begin{aligned}\int_{\Omega_1\times\Omega_2}fd\mu_1\times\mu_2&=\int_{\Omega_1}\left(\int_{\Omega_2}f(\omega_1,\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1)\\&=\int_{\Omega_2}\left(\int_{\Omega_1}f(\omega_1,\omega_2)d\mu_1(\omega_1)\right)d\mu_2(\omega_2).\end{aligned}\tag{1}$$

证　由于方法相同, 只证第一等式.

首先, 由 §4.1 定理 3 知 $\int_{\Omega_2}f(\omega_1,\omega_2)d\mu_2(\omega_2)$ 是 $(\Omega_1,\mathscr{A}_1)$ 上的非负可测函数, 因而等式右端的积分永远存在.

其次, 我们设 $f=\chi_A,A\in\mathscr{A}_1\times\mathscr{A}_2$, 则由乘积测度的定义知

$$\begin{aligned}\int_{\Omega_1\times\Omega_2}fd\mu_1\times\mu_2=\mu_1\times\mu_2(A)&=\int_{\Omega_1}\mu_2(A(\omega_1))d\mu_1(\omega_1)\\&=\int_{\Omega_1}\left(\int_{\Omega_2}\chi_{A(\omega_1)}(\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1)\\&=\int_{\Omega_1}\left(\int_{\Omega_2}\chi_{A(\omega_1,\omega_2)}d\mu_2(\omega_2)\right)d\mu_1(\omega_1)\\&=\int_{\Omega_1}\left(\int_{\Omega_2}f(\omega_1,\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1).\end{aligned}$$

因此当 f 为 $\mathscr{A}_1\times\mathscr{A}_2$ 中集合的示性函数时 (1) 式成立. 由于积分的线性知 f 为非负简单函数时 (1) 式也成立, 设 f 是非负可测函数, 则有 $f_n\uparrow f,\{f_n\}$ 是非负简单

函数序列. 于是由单调收敛定理及非降函数积分的非降性及 (1) 式对非负简单函数成立的事实知

$$\begin{aligned}
&\int_{\Omega_1\times\Omega_2} f d\mu_1\times\mu_2 = \lim_{n\to\infty}\int_{\Omega_1\times\Omega_2} f_n d\mu_1\times\mu_2\\
&=\lim_{n\to\infty}\int_{\Omega_1}\left(\int_{\Omega_2} f_n(\omega_1,\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1)\\
&=\int_{\Omega_1}\lim_{n\to\infty}\left(\int_{\Omega_2} f_n(\omega_1,\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1)\\
&=\int_{\Omega_1}\left(\int_{\Omega_2}\lim_{n\to\infty} f_n(\omega_1,\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1)\\
&=\int_{\Omega_1}\left(\int_{\Omega_2} f(\omega_1,\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1). \qquad \square
\end{aligned}$$

定理 2 设 $(\Omega_i,\mathscr{A}_i,\mu_i), i=1,2$ 是 σ-有限测度空间, f 是 $\mathscr{A}_1\times\mathscr{A}_2$ 可测函数且关于 $\mu_1\times\mu_2$ 的积分存在, 则

i) $g(\omega_1)=\int_{\Omega_2} f(\omega_1,\omega_2)d\mu_2(\omega_2)$, a.e.$(\mu_1)$ 存在且 $\mathscr{A}_1$ 可测;

$h(\omega_2)=\int_{\Omega_1} f(\omega_1,\omega_2)d\mu_1(\omega_1)$, a.e.$(\mu_2)$ 存在且 $\mathscr{A}_2$ 可测;

ii) $\int_{\Omega_1} g d\mu_1, \int_{\Omega_2} h d\mu_2$ 存在且

$$\int_{\Omega_1\times\Omega_2} f d\mu_1\times\mu_2=\int_{\Omega_1} g d\mu_1=\int_{\Omega_2} h d\mu_2;$$

iii) 若 f 关于 $\mu_1\times\mu_2$ 可积, 则 g,h 分别关于 μ_1,μ_2 可积.

证 只证关于 g 的论断. 由于复值函数可以分为实虚两部分考虑, 故只考虑 f 为实可测函数的情形.

i) 由于 $\int_{\Omega_1\times\Omega_2} f d\mu_1\times\mu_2$ 存在, 因而 $\int_{\Omega_1\times\Omega_2} f^+ d\mu_1\times\mu_2$ 或 $\int_{\Omega_1\times\Omega_2} f^- d\mu_1\times\mu_2$ 有限, 不妨设后者有限.

再由定理 1 知

$$\int_{\Omega_1\times\Omega_2} f^\pm d\mu_1\times\mu_2=\int_{\Omega_1}\left(\int_{\Omega_2} f^\pm(\omega_1,\omega_2)\right)d\mu_2 d\mu_1,$$

因而 $\int_{\Omega_2} f^-(\omega_1, \omega_2)d\mu_2$, a.e.$(\mu_1)$ 有限且关于 μ_1 是可积的. 由 §4.1 定理 3 知

$\int_{\Omega_2} f^{\pm}(\omega_1,\omega_2)d\mu_2(\omega_2)$ 是 $\mathscr{A}_1$-可测函数, 所以

$$\begin{aligned} g(\omega_1) &= \int_{\Omega_2} f(\omega_1,\omega_2)d\mu_2(\omega_2) \\ &= \int_{\Omega_2} f^+(\omega_1,\omega_2)d\mu_2(\omega_2) - \int_{\Omega_2} f^-(\omega_1,\omega_2)d\mu_2(\omega_2) \end{aligned}$$

对几乎所有的 ω_1 有意义, 因而是几乎 $\mathscr{A}_1$-可测的. 这就证明了 i).

ii) 由于积分的线性性及 $\int_{\Omega_2} f^-(\omega_1,\omega_2)d\mu_2(\omega_2)$, a.e.$(\mu_1)$ 有限且关于 μ_1 可积, 因而

$$\begin{aligned} \int_{\Omega_1\times\Omega_2} fd\mu_1\times\mu_2 &= \int_{\Omega_1\times\Omega_2} f^+d\mu_1\times\mu_2 - \int_{\Omega_1\times\Omega_2} f^-d\mu_1\times\mu_2 \\ &= \int_{\Omega_1}\left(\int_{\Omega_2} f^+(\omega_1,\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1) \\ &\quad - \int_{\Omega_1}\left(\int_{\Omega_2} f^-(\omega_1,\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1) \\ &= \int_{\Omega_1}\left(\int_{\Omega_2} f^+(\omega_1,\omega_2)d\mu_2(\omega_2)\right. \\ &\qquad \left. - \int_{\Omega_2} f^-(\omega_1,\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1) \\ &= \int_{\Omega_1} g(\omega_1)d\mu_1(\omega_1). \end{aligned}$$

iii) 若 f 关于 $\mu_1\times\mu_2$ 可积, 则 $\int_{\Omega_2} f^{\pm}d\mu_2$ 关于 μ_1 可积, 故 g 关于 μ_1 可积. □

推论 1　设 $(\Omega_i,\mathscr{A}_i,\mu_i), i=1,2$ 是两个 σ-有限测度空间, f 关于 $\mathscr{A}_1\times\mathscr{A}_2$ 可测, 若

$$\int_{\Omega_1}\left(\int_{\Omega_2} |f(\omega_1,\omega_2)|d\mu_2(\omega_2)\right)d\mu_1(\omega_1) < \infty$$

或

$$\int_{\Omega_2}\left(\int_{\Omega_1} |f(\omega_1,\omega_2)|d\mu_1(\omega_1)\right)d\mu_2(\omega_2) < \infty,$$

则 f 关于 $\mu_1\times\mu_2$ 可积, 因而

$$\int_{\Omega_1\times\Omega_2} fd\mu_1\times\mu_2 = \int_{\Omega_1}\left(\int_{\Omega_2} f(\omega_1,\omega_2)d\mu_2(\omega_2)\right)d\mu_1(\omega_1)$$

$$= \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2) d\mu_1(\omega_1) \right) d\mu_2(\omega_2).$$

证 $|f|$ 是 $\mathscr{A}_1 \times \mathscr{A}_2$ 可测函数且非负, 应用定理 1 得

$$\int_{\Omega_1 \times \Omega_2} |f| d\mu_1 \times \mu_2 = \int_{\Omega_1} \left(\int_{\Omega_2} |f(\omega_1, \omega_2)| d\mu_2(\omega_2) \right) d\mu_1(\omega_1)$$

$$= \int_{\Omega_2} \left(\int_{\Omega_1} |f(\omega_1, \omega_2)| d\mu_1(\omega_1) \right) d\mu_2(\omega_2) < \infty,$$

故 f 关于 $\mu_1 \times \mu_2$ 可积, 再应用定理 2 即得所需结论.

注 上面的定理和推论当积分区域是任意可测矩形 $A_1 \times A_2$ 时也都成立. 相应结论的陈述只需将上面定理中的 Ω_1, Ω_2 分别换为 A_1, A_2 即可. 其证明可以对 $f(\omega_1, \omega_2)\chi_{A_1 \times A_2}(\omega_1, \omega_2)$ 应用已得的定理和推论, 并注意到 $\chi_{A_1 \times A_2}(\omega_1, \omega_2) = \chi_{A_1}(\omega_1)\chi_{A_2}(\omega_2)$ 而得出.

一般来说, 对于任意 $G \in \mathscr{A}_1 \times \mathscr{A}_2$, 将上面的定理和推论应用到函数 $f(\omega_1, \omega_2)\chi_G(\omega_1, \omega_2)$, 也可以得出一般区域上的重积分化累次积分的公式. 即

推论 2 若 $(\Omega_i, \mathscr{A}_i, \mu_i)$ 是 σ-有限测度空间, f 是 $\mathscr{A}_1 \times \mathscr{A}_2$ 可测函数,

i) 若 $A_i \in \mathscr{A}_i, i = 1, 2, f\chi_{A_1 \times A_2}$ 满足定理 2 或推论 1 中 f 所应满足的条件, 则

$$\int_{A_1 \times A_2} f d\mu_1 \times \mu_2 = \int_{A_1} \left(\int_{A_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) \right) d\mu_1(\omega_1)$$

$$= \int_{A_2} \left(\int_{A_1} f(\omega_1, \omega_2) d\mu_1(\omega_1) \right) d\mu_2(\omega_2).$$

ii) 若 $G \in \mathscr{A}_1 \times \mathscr{A}_2, f\chi_G$ 满足定理 2 或推论 1 中 f 所满足的条件, 则

$$\int_G f d\mu_1 \times \mu_2 = \int_{\Omega_1} \left(\int_{G(\omega_1)} f(\omega_1, \omega_2) d\mu_2(\omega_2) \right) d\mu_1(\omega_1)$$

$$= \int_{\Omega_2} \left(\int_{G(\omega_2)} f(\omega_1, \omega_2) d\mu_1(\omega_1) \right) d\mu_2(\omega_2).$$

例 1 i) 设 $f_{mn} \geqslant 0, m, n = 1, 2, \cdots$, 则

$$\sum_{m,n=1}^{\infty} f_{mn} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} f_{mn} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} f_{mn}.$$

ii) 设 $\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}|f_{mn}|<\infty$, 则 $\sum_{m,n=1}^{\infty}f_{mn}, \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}f_{mn}, \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}f_{mn}$ 收敛, 且

$$\sum_{m,n=1}^{\infty}f_{mn}=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}f_{mn}=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}f_{mn}.$$

事实上, 设 $\Omega_i=\{1,2,\cdots\}$, $\mathscr{A}_i$ 是 Ω_i 的一切子集作成的 σ-代数, $\mu_i(A)$ 表示集 A 中所含点的个数 (计数测度). 则 $(\Omega_i,\mathscr{A}_i,\mu_i), i=1,2$ 显然是 σ-有限测度空间, $f_{mn}, m,n=1,2,\cdots$ 是 $\mathscr{A}_1\times\mathscr{A}_2$ 可测函数, 故利用定理 1 及定理 2 的推论即可证 i), ii) 成立.

例 2　设 $\{f_n\}$ 是 $(\Omega,\mathscr{A},\mu)$ 上可测函数序列, μ 为 σ-有限测度. 若

$$\sum_{n=1}^{\infty}\int f_n^+d\mu<\infty$$

或

$$\sum_{n=1}^{\infty}\int f_n^-d\mu<\infty,$$

则 $\sum_{n=1}^{\infty}f_n(\omega)$, a.e. 存在, a.e. 可测, 关于 μ 的积分存在且

$$\int_\Omega\sum_{n=1}^{\infty}f_n(\omega)d\mu(\omega)=\sum_{n=1}^{\infty}\int_\Omega f_n(\omega)d\mu(\omega).$$

事实上, 若令 $(\Omega_1,\mathscr{A}_1,\mu_1)=(\Omega,\mathscr{A},\mu)$, 而 $(\Omega_2,\mathscr{A}_2,\mu_2)$ 如例 1 所述. 令 $f(\omega,n)=f_n(\omega), (\omega,n)\in\Omega_1\times\Omega_2$, 则对任一 $B\in\mathscr{B}^{(1)}$,

$$\{(\omega,n):f(\omega,n)\in B\}=\sum_{n=1}^{\infty}\{f_n(\omega)\in B\}\times\{n\}\in\mathscr{A}_1\times\mathscr{A}_2,$$

而 $f^\pm(\omega,n)=f_n^\pm(\omega)$, 对一切 $(\omega,n)\in\Omega_1\times\Omega_2$ 都成立, 由定理 1 易知

$$\begin{aligned}\int_{\Omega_1\times\Omega_2}f^\pm(\omega,n)d\mu_1\times\mu_2&=\int_{\Omega_2}\left(\int_{\Omega_1}f^\pm(\omega,n)d\mu_1(\omega_1)\right)d\mu_2(n)\\&=\sum_{n=1}^{\infty}\int_\Omega f_n^\pm(\omega)d\mu(\omega).\end{aligned}$$

由所给条件知 $\int_{\Omega_1\times\Omega_2}f^+d\mu_1\times\mu_2$ 和 $\int_{\Omega_1\times\Omega_2}f^-d\mu_1\times\mu_2$ 至少有一个有限. 因此 f 关于 $\mu_1\times\mu_2$ 的积分存在, 故定理 2 的条件满足, 因而本例的一切结论都成立.

现在我们来叙述有限维乘积测度空间上重积分化累次积分的定理.

定理 1′ 若 f 是 $(\Omega_1\times\cdots\times\Omega_n,\mathscr{A}_1\times\cdots\times\mathscr{A}_n,\mu_1\times\cdots\times\mu_n)$ 上的非负可测函数,μ_i 是 $(\Omega_i,\mathscr{A}_i)$ 上的 σ-有限测度,$i=1,\cdots,n$ 则对于 $(1,\cdots,n)$ 的任一置换 $(i_1,\cdots,i_n)$, 有

$$\int_{\Omega_1\times\cdots\times\Omega_n} f(\omega_1,\cdots,\omega_n)d\mu_1\times\cdots\times\mu_n$$
$$=\int_{\Omega_{i_n}}\left(\cdots\left(\int_{\Omega_{i_1}} f(\omega_1,\cdots,\omega_n)d\mu_{i_1}(\omega_{i_1})\right)\cdots\right)d\mu_{i_n}(\omega_{i_n}) \tag{2}$$

证 使用函数形式单调类定理来证明 (2) 式, 令 $\mathscr{L}$ 为 $\Omega_1\times\cdots\times\Omega_n$ 上的一切函数所作成的函数类,

$$\mathfrak{M}=\{f\in\mathscr{A}_1\times\cdots\times\mathscr{A}_n,f\geqslant 0\text{且使对}(1,\cdots,n)\text{的任一置换}(i_1,\cdots,i_n)(2)\text{式成立}\},$$

$$\mathscr{C}=\{A_1\times\cdots\times A_n,A_i\in\mathscr{A}_i,i=1,\cdots,n\}$$

$\mathscr{C}$ 是 π 系且 $\sigma(\mathscr{C})=\mathscr{A}_1\times\cdots\times\mathscr{A}_n$, 由积分的单调收敛定理可知 L 为 $\mathscr{L}$-系, 故只需证明当 $f=\chi_{A_1\times\cdots\times A_n}$ 时 (2) 式成立即可, 事实上

$$\int_{\Omega_{i_n}}\left(\cdots\left(\int_{\Omega_{i_1}}\chi_{A_1\times\cdots\times A_n}(\omega_1,\cdots,\omega_n)d\mu_{i_1}(\omega_{i_1})\right)\cdots\right)d\mu_{i_n}(\omega_{i_n})$$
$$=\int_{\Omega_{i_n}}\left(\cdots\left(\int_{\Omega_{i_1}}\chi_{A_1(\omega_1)}\cdots\chi_{A_n(\omega_n)}d\mu_{i_1}(\omega_{i_1})\right)\cdots\right)d\mu_{i_n}(\omega_{i_n})$$
$$=\mu_{i_1}(A_{i_1})\int_{\Omega_{i_n}}\cdots\int_{\Omega_{i_2}}\chi_{A_{i_2}(\omega_{i_2})}\cdots\chi_{A_{i_n}(\omega_{i_n})}d\mu_{i_2}(\omega_{i_2})\cdot d\mu_{i_n}(\omega_{i_n})$$
$$=\mu_{i_1}(A_{i_1})\cdots\mu_{i_n}(A_{i_n})=\mu_1(A_1)\cdots\mu_n(A_n)=\mu_1\times\cdots\times\mu_n(A_1\times\cdots\times A_n)$$
$$=\int_{\Omega_1\times\cdots\times\Omega_n}\chi_{A_1\times\cdots\times A_n}d\mu_1\times\cdots\times\mu_n$$

应该注意本定理的结论包括 (虽然没有明显陈述):

i) $\int_{\Omega_{i_1}} f(\omega_1,\cdots,\omega_n)d\mu_{i_1}(\omega_{i_1})$ 是 $\mathscr{A}_1\times\cdots\times\mathscr{A}_{i_1-1}\times\mathscr{A}_{i_1+1}\times\cdots\times\mathscr{A}_n$ 非负可测函数;

ii) $\int_{\Omega_{i_2}}\left(\int_{\Omega_{i_1}} f(\omega_1,\cdots,\omega_n)d\mu_{i_1}(\omega_{i_1})\right)d\mu_{i_2}(\omega_{i_2})$ 是 $\mathscr{A}_1\times\cdots\times\mathscr{A}_{i_1-1}\times\mathscr{A}_{i_1+1}\times\cdots\times\mathscr{A}_{i_2-1}\times\mathscr{A}_{i_2+1}\times\cdots\times\mathscr{A}_n$ 非负可测函数 (这里 $1\leqslant i_1<i_2\leqslant n$, 事实上当 $i_2<i_1$ 时亦真, 只是 σ-代数表示法不同).

……

定理 2′　设 $(\Omega_i, \mathscr{A}_i, \mu_i), i = 1, 2, \cdots, n$ 是 σ-有限测度空间,f 关于 $\mathscr{A}_1 \times \cdots \times \mathscr{A}_n$ 可测且关于 $\mu_1 \times \cdots \times \mu_n$ 的积分存在, 则对 $(1, \cdots, n)$ 的任一满足条件 $1 \leqslant j_1 < \cdots < j_{n-k} \leqslant n$ 的置换 $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$, 有

i) $g(\omega_{j_1}, \cdots, \omega_{j_{n-k}}) = \displaystyle\int_{\Omega_{i_k}} \left(\cdots \left(\int_{\Omega_{i_1}} f(\omega_1, \cdots, \omega_n) d\mu_{i_1}(\omega_{i_1}) \right) \cdots \right) d\mu_{i_k}(\omega_{i_k})$

除去一个 $\mathscr{A}_{j_1} \times \cdots \times \mathscr{A}_{j_{n-k}}$ 可测的关于测度 $\mu_{j_1} \times \cdots \times \mu_{j_{n-k}}$ 为零的集合外存在且可测;

ii) $\displaystyle\int_{\Omega_{j_1} \times \cdots \times \Omega_{j_{n-k}}} g d\mu_{j_1} \times \cdots \times \mu_{j_{n-k}}$ 存在, 且

$$\int_{\Omega_1 \times \cdots \times \Omega_n} f d\mu_1 \times \cdots \times \mu_n = \int_{\Omega_{j_1} \times \cdots \times \Omega_{j_{n-k}}} g d\mu_{j_1} \times \cdots \times \mu_{j_{n-k}}.$$

特别取 $k = n-1, j_1 = i_n$, 即得 (2), 即定理 1′ 的结论.

iii) 若 f 关于 $\mu_1 \times \cdots \times \mu_n$ 可积, 则 g 关于 $\mu_{j_1} \times \cdots \times \mu_{j_{n-k}}$ 可积.

定理 2′ 可以仿照定理 2 应用定理 1′ 来证明.

推论 1′　设 f 是 $(\Omega_1 \times \cdots \times \Omega_n, \mathscr{A}_1 \times \cdots \times \mathscr{A}_n, \mu_1 \times \cdots \times \mu_n)$ 上的可测函数, 如果对于 $(1, \cdots, n)$ 的某一置换 $(i_1, \cdots, i_n)$ 来说

$$\int_{\Omega_{i_n}} \left(\cdots \left(\int_{\Omega_{i_1}} |f(\omega_1, \cdots, \omega_n)| d\mu_{i_1}(\omega_{i_1}) \right) \cdots \right) d\mu_{i_n}(\omega_{i_n}) < \infty,$$

则 f 关于 $\mu_1 \times \cdots \times \mu_n$ 可积, 因而 (2) 式成立.

推论 2′　设 f 是 $(\Omega_1 \times \cdots \times \Omega_n, \mathscr{A}_1 \times \cdots \times \mathscr{A}_n, \mu_1 \times \cdots \times \mu_n)$ 上的可测函数,

i) 若 $A_i \in \mathscr{A}_i, i = 1, \cdots, n, f\chi_{A_1 \times \cdots \times A_n}$ 满足定理 2′ 或推论 1′ 中 f 所应满足的条件, 则

$$\begin{aligned}
&\int_{A_1 \times \cdots \times A_n} f d\mu_1 \times \cdots \times \mu_n \\
&= \int_{A_{i_n}} \left(\cdots \left(\int_{A_{i_1}} f(\omega_1, \cdots, \omega_n) d\mu_{i_1}(\omega_{i_1}) \right) \cdots \right) d\mu_{i_n}(\omega_{i_n}),
\end{aligned}$$

其中 $(i_1 \cdots, i_n)$ 是 $(1, \cdots, n)$ 的任一置换.

ii) 若 $G \in \mathscr{A}_1 \times \cdots \times \mathscr{A}_n, f\chi_G$ 满足定理 2′ 或推论 1′ 中 f 所应满足的条件, 则

$$\begin{aligned}
&\int_G f d\mu_1 \times \cdots \times \mu_n \\
&= \int_{\Omega_{i_1} \times \cdots \times \Omega_{i_k}} \int_{G(\omega_{i_1}, \cdots, \omega_{i_k})} f(\omega_1, \cdots, \omega_n) d\mu_{i_{k+1}} \times \cdots \times \mu_{i_n} d\mu_{i_1} \times \cdots \times \mu_{i_k}
\end{aligned}$$

其中 $(i_1, \cdots, i_n)$ 是 $(1, \cdots, n)$ 的任一置换, $i_1 < \cdots < i_k$, 而 $G(\omega_{i_1}, \cdots, \omega_{i_k})$

$$= \left\{ (\omega_{j_1}, \cdots, \omega_{j_{n-k}}): \begin{array}{l} (\omega_1, \cdots, \omega_n) \in G, j_1 < \cdots < j_{n-k}, \\ j_l \neq s, l = 1, \cdots, n-k, s = i_1, \cdots, i_k. \end{array} \right\}$$

推论 3 如果 f_i 是 $(\Omega_i, \mathscr{A}_i, \mu_i)$ 上的可积函数,$i = 1, \cdots, n$, 则 $f(\omega_1, \cdots, \omega_n) = f_1(\omega_1) \cdots f_n(\omega_n)$ 是 $(\Omega_1 \times \cdots \times \Omega_n, \mathscr{A}_1 \times \cdots \times \mathscr{A}_n, \mu_1 \times \cdots \times \mu_n)$ 上的可积函数, 且对任何 $A_i \in \mathscr{A}_i, i = 1, \cdots, n$ 有

$$\int_{A_1 \times \cdots \times A_n} f d\mu_1 \times \cdots \times \mu_n = \int_{A_1} f_1(\omega_1) d\mu_1 \cdots \int_{A_n} f_n(\omega_n) d\mu_n. \tag{3}$$

这些推论的证明不在此赘述了，请读者作为练习来完成.

上述结果在解决概率论的一些理论问题上很有用，特别是对于连续型随机变量，有些问题的彻底解决有赖于 Fubini 定理.

首先有

$$R^{(m_1+\cdots+m_n)} = R^{(m_1)} \times \cdots \times R^{(m_n)},$$

$$\mathscr{B}^{(m_1+\cdots+m_n)} = \mathscr{B}^{(m_1)} \times \cdots \times \mathscr{B}^{(m_n)},$$

$$\Lambda^{(m_1+\cdots+m_n)} = \Lambda^{(m_1)} \times \cdots \times \Lambda^{(m_n)},$$

其中 $R^{(m)}$ 是 m 维实欧氏空间，$\mathscr{B}^{(m)}$ 是 m 维 Borel 域，$\Lambda^{(m)}$ 是 $\mathscr{B}^{(m)}$ 上的 L-测度.

下面我们首先讨论连续型随机变量独立的判定定理.

定理 3 若 $\xi^{(k)} = (\xi_{k1}, \cdots, \xi_{km_k}), k = 1, \cdots, n$ 是具有分布密度 $p_k(x^{(k)})$ 的 n 个独立的多维随机变量, 则 $(\xi_{11}, \cdots, \xi_{nm_n})$ (简记作 $(\xi^{(1)}), \cdots, \xi^{(n)})$ 为连续型随机变量, 其分布密度

$$p(x^{(1)}, \cdots, x^{(n)}) = p_1(x^{(1)}) \cdots p_n(x^{(n)}), x^{(k)} \in R^{(m_k)},$$

$$k = 1, \cdots, n. \tag{4}$$

上式除去一个 $\Lambda^{(m_1+\cdots+m_n)}$ 零测集而外处处成立.

反之, 若 $(\xi^{(1)}, \cdots, \xi^{(n)})$ 是连续型随机变量, 其分布密度为 $p(x^{(1)}, \cdots, x^{(n)})$, 且存在 n 个 Borel 可测函数 $q_k(x^{(k)}), x^{(k)} \in R^{(m_k)}, k = 1, \cdots, n$ 使

$$p(x^{(1)}, \cdots, x^{(n)}) = q_1(x^{(1)}) \cdots q_n(x^{(n)}), \text{a.e.}(\Lambda^{(m_1+\cdots+m_n)}), \tag{5}$$

则 $\xi^{(k)}, k = 1, \cdots, n$ 独立, 且 $\int q_k(x^{(k)}) d\Lambda^{(m_k)} \neq 0, k = 1, \cdots, n. \xi^{(k)}$ 具有分布密度

$$p_{\xi^{(k)}}(x^{(k)}) = \frac{q_k(x^{(k)})}{\int q_k(x^{(k)}) d\Lambda^{(m_k)}}, x^{(k)} \in R^{(m_k)}, k = 1, \cdots, n.$$

证　i) 先证定理的第一部分. 由于 $\xi^{(k)}, k=1,\cdots,n$ 独立, 且各具有分布密度 $p_k(x^{(k)})$. 设 $\xi^{(k)}$ 的分布函数是 $F_k(x^{(k)})$, 由独立随机变量的定义知 $(\xi^{(1)},\cdots,\xi^{(n)})$ 的分布函数

$$F(x^{(1)},\cdots,x^{(n)})=\prod_{k=1}^{n}F_k(x^{(k)})=\prod_{k=1}^{n}\int_{(-\infty,x^{(k)})}p_k(t^{(k)})d\Lambda^{(m_k)},$$

其中 $x^{(k)},t^{(k)}\in R^{(m_k)}$, 再由推论 3 知 $\prod_{k=1}^{n}p_k(t^{(k)})$ 是非负可积的, 且

$$F(x^{(1)},\cdots,x^{(n)})=\int_{(-\infty,x^{(1)})\times\cdots\times(-\infty,x^{(n)})}\prod_{k=1}^{n}p_k(t^{(k)})d\Lambda^{(m_1+\cdots+m_n)}.$$

这就是说$(\xi^{(1)},\cdots,\xi^{(n)})$=$(\xi_{11},\cdots\xi_{nm_n})$是连续型随机变量, 具有分布密度$\prod_{k=1}^{n}p_k(x^{(k)})$.

ii) 再证定理的第二部分. 由于 $q_k(x^{(k)}), k=1,\cdots,n$ 是 Borel 可测函数且 $q_1(x^{(1)})\cdots q_n(x^{(n)})=p(x^{(1)},\cdots,x^{(n)})$, a.e. 因而每一 $q_k(x^{(k)})$ 是 a.e. 有限的, $q_1(x^{(1)})\cdots q_n(x^{(n)})$ 是 a.e. 可测并且关于 $\Lambda^{(m_1+\cdots+m_n)}$ 可积的. 由推论 1′ 知 (2) 式成立. 对一切 $x^{(k)}\in R^{(m_k)}, k=1,\cdots,n$

$$\begin{aligned}F(x^{(1)},\cdots,x^{(n)})&=\int_{\prod_{k=1}^{n}(-\infty,x^{(k)})}p(t^{(1)},\cdots,t^{(n)})d\Lambda^{(m_1+\cdots+m_n)}\\&=\int_{\prod_{k=1}^{n}(-\infty,x^{(k)})}q_1(t^{(1)})\cdots q_n(t^{(n)})d\Lambda^{(m_1+\cdots+m_n)}\\&=\int_{(-\infty,x^{(n)})}\cdots\left(\int_{(-\infty,x^{(1)})}q_1(t^{(1)})\cdots q_n(t^{(n)})d\Lambda^{(m_1)}\right)\cdots d\Lambda^{(m_n)}\\&=\prod_{k=1}^{n}\int_{(-\infty,x^{(k)})}q_k(t^{(k)})d\Lambda^{(m_k)},\end{aligned}$$

令

$$G_k(x^{(k)})=\int_{(-\infty,x^{(k)})}q_k(t^{(k)})d\Lambda^{(m_k)},$$

则由第 2 章 §2.4 推论 4 知 $G_k(\infty)\neq 0$, 且 $\xi^{(k)}, k=1,\cdots,n$, 独立, $\xi^{(k)}$ 的分布函数

为 $G_k(x^{(k)})/G_k(\infty)$, 即

$$F_{\xi^{(k)}}(x^{(k)})=\int_{(-\infty,x^{(k)})}q_k(t^{(k)})d\Lambda^{(m_k)}\Big/\int_{R^{(m_k)}}q_k(t^{(k)})d\Lambda^{(m_k)},$$

$$x^{(k)}\in R^{(m_k)}.$$

因而 $\xi^{(k)}$ 的分布密度为

$$p_{\xi^{(k)}}(x^{(k)})=q_k(t^{(k)})\Big/\int_{R^{(m_k)}}q_k(t^{(k)})d\Lambda^{(m_k)},\ \text{a.e.}(\Lambda^{(m_k)}).$$ □

定理 3 的第二部分对于验证随机变量的独立性带来很大的方便.

例 3 设 ξ_1,ξ_2 独立且均按 $N(0,\sigma^2)$ 分布,$\eta_1=\sqrt{\xi_1^2+\xi_2^2},\eta_2=\xi_1/\xi_2$, 则 η_1 与 η_2 独立.

事实上, 应用第 3 章 §3.5 定理 5 可以求得 η_1,η_2 是连续型随机变量且其联合分布密度

$$p_{\eta_1,\eta_2}(y_1,y_2)=\frac{1}{\pi\sigma^2}y_1\mathrm{e}^{-\frac{y_1^2}{2\sigma^2}}\chi_{[0,\infty)}(y_1)\frac{1}{1+y_2^2},(y_1,y_2)\in R^{(2)}$$

(参看 [3]85 页例 5) 而 $\frac{1}{\pi\sigma^2}y_1\mathrm{e}^{-\frac{y_1^2}{2\sigma^2}}\chi_{[0,\infty)}(y_1)$ 和 $\frac{1}{1+y_2^2}$ 都是一元 Borel 可测函数, 因而 η_1 与 η_2 独立. (注意: 这里并不需要验证 $\int_0^\infty\frac{y_1}{\sigma^2}\mathrm{e}^{-\frac{y_1}{2\sigma^2}}\mathrm{d}y_1=1$ 和 $\int\frac{1}{\pi(1+y_2^2)}$ $\mathrm{d}y_2=1$, 也不需要求出 η_1 与 η_2 的分布密度, 即可断定 η_1 与 η_2 是独立的.)

例 4 设 $\xi=(\xi_1,\cdots,\xi_n)$ 是按 $N(m,D)$ 分布的随机变量, 则存在 ξ 的一个线性变换,$\eta=\xi\cdot C'+b$, 其中 C 是 $n\times n$ 正交方阵,$b=-m\cdot C'$ 使得 η_k 按 $N(0,\sigma_k^2)$ 分布, 且 $\eta_1,\cdots,\eta_n$ 独立.

事实上, 由于 D 是 $n\times n$ 定正方阵, 由线性代数知有一 $n\times n$ 正交方阵 C 存在使得

$$CDC'=\Lambda=\begin{pmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_n^2\end{pmatrix},\sigma_k>0.$$

于是由 $CC'=I$ 及矩阵运算知 $|D|=\sigma_1^2\cdots\sigma_n^2$, 而

$$D^{-1}=C'\Lambda^{-1}C.$$

令 $\eta=\xi C'-mC'$, 由第 3 章 §3.5 例 4 及 $\|C\|=1$ 知

$$\begin{aligned}p_\eta(y_1,\cdots,y_n)&=p_\xi((y+mC')C)\\&=\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|D|}}\times\exp\left\{-\frac{1}{2}[(y+mC')C-m]D^{-1}[(y+mC')C-m]'\right\}\\&=\frac{1}{(2\pi)^{\frac{n}{2}}\sigma_1\cdots\sigma_n}\exp\left\{-\frac{1}{2}yCD^{-1}C'y'\right\}\\&=\frac{1}{(2\pi)^{\frac{n}{2}}\sigma_1\cdots\sigma_n}\exp\left\{-\frac{1}{2}y\varLambda^{-1}y'\right\}\\&=\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi}\sigma_k}\exp\left\{-\frac{y_k^2}{2\sigma_k^2}\right\}.\end{aligned}$$

故 η_k 按 $N(0,\sigma_k^2)$ 分布, 且 $\eta_1\cdots,\eta_n$ 独立.

Fubini 定理在概率论中的另一个应用是当已知一连续型随机变量 $(\xi_1,\cdots,\xi_n)$ 的分布密度为 $p(x_1,\cdots,x_n)$ 时, 则由它的一部分分量作成的随机变量 $(\xi_{i_1},\cdots,\xi_{i_k})$ 的分布密度可以由 $p(x_1,\cdots,x_n)$ 的积分计算出来.

定理 4　如果 $(\xi_1,\cdots,\xi_n)$ 是连续型随机变量, 其分布密度为 $p(x_1,\cdots,x_n)$, 则对于任意的 $1\leqslant i_1<\cdots<i_k\leqslant n$ 来说, $(\xi_{i_1},\cdots,\xi_{i_k})$ 也是一连续型随机变量且它的分布密度

$$p_{\xi_{i_1},\cdots,\xi_{i_k}}(x_{i_1},\cdots,x_{i_k})=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(x_1,\cdots,x_n)dx_{j_1}\cdots dx_{j_{n-k}},$$

其中 $(i_1,\cdots,i_k,j_1,\cdots,j_{n-k})$ 是 $(1,\cdots,n)$ 的任一个 (满足 $1\leqslant i_1<\cdots<i_k\leqslant n$) 的置换.

证　我们只证 $(\xi_1,\cdots,\xi_k)$ 是连续型随机变量且分布密度为

$$p_{\xi_1,\cdots,\xi_k}(x_1,\cdots,x_k)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(x_1,\cdots,x_n)dx_{k+1}\cdots dx_n.$$

一般情形证法类似, 只是叙述更加复杂一些.

设 $F(x_1,\cdots,x_n)$ 是 $(\xi_1,\cdots,\xi_n)$ 的分布函数, 我们来计算 $(\xi_1,\cdots,\xi_k)$ 的分布函数. 由分布函数的性质 (第 2 章 2.3.1 节性质 4 的 (15) 式及其证明) 及 Fubini 定

理, 对任意的 $(x_1,\cdots,x_k)\in R^{(k)}$.

$$\begin{aligned}
&F_{\xi_1,\cdots,\xi_k}(x_1,\cdots,x_k)\\
&=p\{(\xi_1,\cdots,\xi_k)\in(-\infty,x_1)\times\cdots\times(-\infty,x_k)\}\\
&=p\{(\xi_1,\cdots,\xi_n)\in(-\infty,x_1)\times\cdots\times(-\infty,x_k)\\
&\quad\times(-\infty,\infty)\times\cdots\times(-\infty,\infty)\}\\
&=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_k}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(t_1,\cdots,t_n)dt_1\cdots dt_n\\
&=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_k}\left[\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(t_1,\cdots,t_n)dt_{k+1}\cdots dt_n\right]dt_1\cdots dt_k,
\end{aligned}$$

其中 $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(t_1,\cdots,t_n)dt_{k+1}\cdots dt_n$ 是 $(R^{(k)},\mathscr{B}^{(k)})$ 上非负可测函数, 因而 $(\xi_1,\cdots,\xi_k)$ 是连续型随机变量且它的分布密度

$$p_{(\xi_1,\cdots,\xi_k)}(x_1,\cdots,x_k)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(x_1,\cdots,x_n)dx_{k+1}\cdots dx_n. \qquad \square$$

第 2 章 2.3.1 节例 1 的计算 (由二维正态分布求边缘分布) 实际上是根据这一定理来进行的. 第 3 章 3.5.3 节例 5 在计算两个连续型随机变量乘积的分布律时也应用了这一定理.

最后, 我们再应用乘积测度与 Fubini 定理求出两个多维独立随机变量和的分布律, 并引出两个测度及两个概率密度的卷积的概念.

定理 5 设 ξ,η 都是 n 维随机变量, 且 ξ,η 独立, 分别具有分布律 P_ξ,P_η, 则 $\zeta=\xi+\eta$ 的分布律为

$$P_\zeta(B)=\int P_\eta(B-x)dP_\xi(x)=\int P_\xi(B-y)dP_\eta(y), B\in\mathscr{B}^{(n)},$$

其中 $B-x=\{z-x,z\in B\}$.

证 由于 ξ,η 独立, 所以 (ξ,η) 的分布律为 $P_\xi\times P_\eta$, 根据第 3 章 §3.4 所述随机变量函数的分布律的 L-S 积分表示, 对任意 $B\in\mathscr{B}^{(n)}$ 有

$$\begin{aligned}
P_\zeta(B)&=P_\xi\times P_\eta(\{(x,y):x+y\in B,x,y\in R^{(n)}\})\\
&=\int_{x+y\in B}dP_\xi\times P_\eta=\int_{R^{(2n)}}\chi_B(x+y)dP_\xi\times P_\eta.
\end{aligned}$$

根据 Fubini 定理, 化重积分为累次积分,

$$P_\xi(B)=\int_{R^{(n)}}\left[\int_{R^{(n)}}\chi_{B-x}(y)dP_\eta(y)\right]dP_\xi(x)$$
$$=\int_{R^{(n)}}P_\eta(B-x)dP_\xi(x).$$

同理可证

$$P_\zeta(B)=\int_{R^{(n)}}P_\xi(B-y)dP_\eta(y).$$ □

定义　设 λ,μ 是 $(R^{(n)},\mathscr{B}^{(n)})$ 上的测度, 称测度

$$(\lambda*\mu)(B)=\int\lambda(B-y)d\mu(y)$$
$$=\int\mu(B-x)d\lambda(x)=(\mu*\lambda)(B),\quad B\in\mathscr{B}^{(n)}$$

为 λ 和 μ 的卷积.

这样, 我们可以说二独立随机变量和的分布律是二独立随机变量分布律的卷积.

对于连续型随机变量有下列的

推论　设 ξ,η 是分别具有分布密度 $p_\xi(x)$, $p_\eta(x)$ 的独立 n 维随机变量, 则 $\zeta=\xi+\eta$ 为连续型随机变量, 其分布密度为

$$p_\zeta(x)=\int_{-\infty}^{\infty}p_\xi(x-y)p_\eta(y)dy$$
$$=\int_{-\infty}^{\infty}p_\eta(x-y)p_\xi(y)dy=p_\xi*p_\eta(x),$$

$p_\xi*p_\eta$ 是 $p_\xi(x)$ 和 $p_\eta(x)$ 的卷积函数.

习题及补充

1. 试问: 以 $\overline{\mathscr{A}_1\times\mathscr{A}_2}$ 代替 Fubini 定理中的 $\mathscr{A}_1\times\mathscr{A}_2$, 该定理是否还成立? (提示: 应用第 2 章 §2.2 习题 12 及 a.e. 可测函数积分的定义.)

2. 设 $f_i(\omega_i)$ 是 σ-有限测度空间 $(\Omega_i,\mathscr{A}_i,\mu_i), i=1,2$ 上的可积函数, 试证:

i) $g_i(\omega_1,\omega_2)=f_i(\omega_i), i=1,2$, 是 $\mathscr{A}_1\times\mathscr{A}_2$ 可测的;

ii) $h(\omega_1,\omega_2)=f_1(\omega_1)f_2(\omega_2)$, 是 a.e. $\mathscr{A}_1\times\mathscr{A}_2$ 可测的;

iii) h 关于 $\mu_1\times\mu_2$ 可积且

$$\int_{\Omega_1\times\Omega_2}hd\mu_1\times\mu_2=\int_{\Omega_1}f_1d\mu_1\cdot\int_{\Omega_2}f_2d\mu_2.$$

3. 设 μ_1,μ_2 是 σ-有限测度,$f(\omega_1,\omega_2)$ 是 $(\Omega_1\times\Omega_2,\mathscr{A}_1\times\mathscr{A}_2,\mu_1\times\mu_2)$ 上平方可积函数, $h(\omega_2)$ 是 $(\Omega_2,\mathscr{A}_2,\mu_2)$ 上平方可积函数. 试证 $\int_{\Omega_2}f(\omega_1,\omega_2)h(\omega_2)d\mu_2(\omega_2)$ 是 $(\Omega_1,\mathscr{A}_1,\mu_1)$ 上平方可积函数.

4. 证明定理 2′

5. 设 $f(\omega_1,\omega_2)$ 是 $(\Omega_1\times\Omega_2,\mathscr{A}_1\times\mathscr{A}_2,\mu_1\times\mu_2)$ 上的可测函数. 则

$$f(\omega_1,\omega_2)=0,\quad \text{a.e.}(\mu_1\times\mu_2)$$

当且仅当存在一个 $N_1\in\mathscr{A}_1,\mu_1(N_1)=0$, 对一切 $\omega_1\in N_1^c$, 存在 N_2(与 ω_1 有关)$\in\mathscr{A}_2,\mu_2(N_2)=0$, 当 $\omega_2\in N_2^c$ 时 $f(\omega_1,\omega_2)=0$.

6. 利用 Fubini 定理证明独立随机变量数学期望的乘法定理 (第 3 章 3.4.1 性质 3).

7. 设 ξ_1,ξ_2 是连续型独立随机变量,$p_1(x),p_2(x)$ 是它们的分布密度, 试应用乘积概率及 Fubini 定理证明 $\xi_1+\xi_2,\xi_1-\xi_2,\xi_1\cdot\xi_2$ 及 ξ_1/ξ_2 (此时需要假设 $p_2(x)>0$, a.e. 对 L-测度) 是连续型随机变量, 并求出它们的分布密度.

8. 设 $\xi_1,\cdots,\xi_n$ 独立且均按 $N(a,\sigma^2)$ 分布. 试证

$$\bar{x}=\frac{1}{n}(\xi_1+\cdots+\xi_n)$$

与

$$s^2=\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{x})^2$$

独立.

9. 对于 $(R^{(n)},\mathscr{B}^{(n)})$ 上的三个概率测度 P_1,P_2 和 P_3, 试证

$$(P_1*P_2)*P_3=P_1*(P_2*P_3).$$

若 P_1,P_2 和 P_3 分别是 ξ,η,ζ 的分布律且 ξ,η,ζ 独立, 试证 $\xi+\eta+\zeta$ 的分布律是 $(P_1*P_2)*P_3$.

10. 设 $\zeta=\xi+\eta$ 具有退化分布 (即存在 $x\in R^{(n)}$ 使 $P(\zeta=x)=1$) 而 ξ,η 独立, 则 ξ,η 必具有退化分布, 且若 $P(\xi=y)=1,P(\eta=z)=1$, 则 $x=y+z$.

11. 设 (ξ_1,ξ_2) 的分布密度 $p(x_1,x_2)$ 是连续函数, 则 ξ_1,ξ_2 独立的充分必要条件是存在连续函数 $p_i(x_i),i=1,2$, 使得 $p(x_1,x_2)=p_1(x_1)p_2(x_2),(x_1,x_2)\in R^{(2)}$.

12. 设 ξ,η 独立,$r\geqslant 1$, 若 $E\xi=E\eta=0$, 则

$$E|\xi|^r\leqslant E|\xi+\eta|^r,E|\eta|^r\leqslant E|\xi+\eta|.$$

13. 在定理 2 中, 如果不假定 $\int fd\mu_1\times\mu_2$ 存在,而代之以假定 $\int\left[\int fd\mu_2\right]d\mu_1$

存在, 则定理是不成立的. 为证明这点, 作为例子考虑一个正的不可积的随机变量 ξ, 它定义在 $(\Omega_1, \mathscr{A}_1, P_1)$ 上, 并在此空间和具有相等概率的两点离散空间 $\{0,1\}$ 的乘积空间上定义随机变量 η

$$\eta(\omega, 0) = \xi(\omega),$$

$$\eta(\omega, 1) = -\xi(\omega).$$

14. 设 λ, μ 是 $(R^{(n)}, \mathscr{B}^{(n)})$ 上有限 L-S 测度, F_1, F_2 是其相应的分布函数. 试证: $\lambda * \mu$ 的分布函数是

$$F(x) = \int_{-\infty}^{\infty} F_1(x-y) dF_2(y).$$

§4.3　无穷乘积概率空间

在引言中我们提到需要构造独立随机变量序列的概率场. 在随机过程论中需要对任意参数集 T 构造一个概率场, 使得在这个概率场上 $\{\xi_t, t \in T\}$ 是独立随机变量族. 这就提出构造无穷乘积概率空间的问题.

设 T 是任意给定的无限个元素的参数集, 常用的是 $T = \{1, 2, \cdots\}, T = [a, b]$ 或 $T = (a, b), a, b \in R^{(1)}$ 或 $\tilde{R}^{(1)}$ 等等. 对每个 $t \in T, (\Omega_t, \mathscr{A}_t, P_t)$ 是概率空间. 自然我们定义 $\Omega_t, t \in T$ 的乘积空间为

$$\prod_{t \in T} \Omega_t = \{\omega_T : \omega_T = \{\omega(t), t \in T\}, \omega(t) \in \Omega_t, t \in T\}. \tag{1}$$

但在 $\prod\limits_{t \in T} \Omega_t$ 中构造乘积 σ-域和独立乘积测度就不是很明显的了. 我们将会看到 $P_t(\Omega_t) = 1, t \in T$ 在独立乘积测度的构造中占有特别重要的地位, 不是一个普通的条件, 也就是说, 本质上不存在一般的无穷乘积测度空间, 在这一节我们研究无穷乘积概率空间的构造.

在构造无穷乘积 σ-代数时, 要借助于有限维的结果. 我们把形如

$$\begin{aligned} &B^{T_N} \times \prod_{t \in T - T_N} \Omega_t : T_N = \{t_1, \cdots, t_N\} \subset T, N\text{有限}, \\ &B^{T_N} \in \mathscr{A}_{t_1} \times \cdots \times \mathscr{A}_{t_N} \end{aligned}$$

的集称为一个**可测柱集**, B^{T_N} 称为该**柱集的底**,令

$$\begin{aligned} \mathscr{D}^T &= \{\text{一切可测柱集全体}\} \\ &= \left\{ B^{T_N} \times \prod_{t \in T - T_N} \Omega_t : \begin{array}{l} T_N = \{t_1, \cdots, t_N\} \subset T, N\text{有限}, \\ B^{T_N} \in \mathscr{A}_{t_1} \times \cdots \times \mathscr{A}_{t_N} \end{array} \right\}. \end{aligned} \tag{2}$$

我们把 $\mathscr{D}^T$ 上的最小 σ-代数 $\sigma(\mathscr{D}^T)$ 称为 $\{\mathscr{A}_t, t\in T\}$ 的乘积 σ-代数, 记作 $\prod\limits_{t\in T}\mathscr{A}_t$.

我们首先构造可数无穷乘积概率. 为此引入以下记号:

$$\Omega \triangleq \Omega_1\times\Omega_2\times\cdots=\prod_{i=1}^{\infty}\Omega_i,$$

$$\Omega^{(n)}\triangleq\Omega_{n+1}\times\Omega_{n+2}\times\cdots=\prod_{i=n+1}^{\infty}\Omega_i, \tag{3}$$

$$\mathscr{C}\triangleq\{A_1\times\cdots\times A_n\times\Omega^{(n)}, A_i\in\mathscr{A}_i, i=1,\cdots,n, n\text{为任意正整数}\}, \tag{4}$$

$$\mathscr{D}\triangleq\{A\times\Omega^{(n)}: A\in\mathscr{A}_1\times\cdots\times\mathscr{A}_n, n\text{为任意正整数}\}, \tag{5}$$

$$\mathscr{D}^{(n)}\triangleq\{A\times\Omega^{(n+m)}: A\in\mathscr{A}_{n+1}\times\cdots\times\mathscr{A}_{n+m}, m\text{为任意正整数}\}, \tag{6}$$

$$\mathscr{A}\triangleq\mathscr{A}_1\times\mathscr{A}_2\times\cdots=\prod_{n=1}^{\infty}\mathscr{A}_n, \tag{7}$$

$$\mathscr{A}^{(n)}\triangleq\mathscr{A}_{n+1}\times\mathscr{A}_{n+2}\times\cdots=\prod_{m=1}^{\infty}\mathscr{A}_{n+m}. \tag{8}$$

容易证明 $\mathscr{C}$ 是 Ω 上的半集代数, $\mathscr{D}$ 是 Ω 上的集代数且 $\mathscr{A}=\sigma(\mathscr{D})=\sigma(\mathscr{C})$, 而 $\mathscr{D}^{(n)}$ 是 $\Omega^{(n)}$ 上的集代数且 $\mathscr{A}^{(n)}=\sigma(\mathscr{D}^{(n)})$.

对于截集的概念可以对照有限维来定义. 容易证明，若 $A\in\mathscr{D}, A=A'\times\Omega^{(m)}, A'\in\mathscr{A}_1\times\cdots\times\mathscr{A}_n\times\cdots\times\mathscr{A}_m, m>n$, 则 $A(\omega_1,\cdots,\omega_n)\in\mathscr{D}^{(n)}, A'(\omega_1,\cdots,\omega_n)\in\mathscr{A}_{n+1}\times\cdots\times\mathscr{A}_m$, 且 $A(\omega_1,\cdots,\omega_n)=A'(\omega_1,\cdots,\omega_n)\times\Omega^{(m)}$.

现在我们来证明可数维无穷乘积概率定理.

定理 1 设 $(\Omega_i,\mathscr{A}_i,P_i), i=1,2,\cdots$ 是给定的可数无穷个概率空间, 则在 $(\Omega,\mathscr{A})$ 上有唯一的概率满足条件

$$P(A_1\times\cdots\times A_n\times\Omega^{(n)})=P_1(A_1)\cdots P_n(A_n), \tag{9}$$

其中 n 是任意自然数, $A_i\in\mathscr{A}_i, i=1,\cdots,n$. 称 P 为 $P_i, i=1,2,\cdots$ 的无穷乘积概率, 记作 $\prod\limits_{i=1}^{\infty}P_i$. 称 $\left(\prod\limits_{i=1}^{\infty}\Omega_i,\prod\limits_{i=1}^{\infty}\mathscr{A}_i,\prod\limits_{i=1}^{\infty}P_i\right)$ 为 $(\Omega_i,\mathscr{A}_i,P_i), i=1,2,\cdots$ 的乘积概率空间.

证 我们分三个步骤来证明定理.

(一) 在 $\mathscr{D}$ 上定义一个满足 (9) 的有限可加非负集函数 P, 且 $P(\Omega)=1$.

对任意的 $D\in\mathscr{D}$, 存在 n 及 $A\in\mathscr{A}_1\times\cdots\times\mathscr{A}_n$, 使 $D=A\times\Omega^{(n)}$, 令

$$P(D)=(P_1\times\cdots\times P_n)(A). \tag{10}$$

(i) 如上定义的 P 是 $\mathscr{D}$ 上的集函数, 即 $P(D)$ 与 D 的表示法无关. 事实上, 若

$$D=A\times\Omega^{(n)}=B\times\Omega^{(m)}, n<m,$$

则必有

$$A\times\Omega^{(n)}=(A\times\Omega_{n+1}\times\cdots\times\Omega_m)\times\Omega^{(m)},$$

即

$$A\times\Omega_{n+1}\times\cdots\times\Omega_m=B.$$

于是

$$\begin{aligned}(P_1\times\cdots\times P_m)(B)&=(P_1\times\cdots\times P_m)(A\times\Omega_{n+1}\times\cdots\times\Omega_m)\\&=((P_1\times\cdots\times P_n)\times(P_{n+1}\times\cdots\times P_m))(A\times(\Omega_{n+1}\times\cdots\times\Omega_m))\\&=(P_1\times\cdots\times P_n)(A)(P_{n+1}\times\cdots\times P_m)(\Omega_{n+1}\times\cdots\times\Omega_m)\\&=(P_1\times\cdots\times P_n)(A)\end{aligned}$$

(注意: 这里用到了 $P_i(\Omega_i)=1, i=1,2,\cdots$). 这就证明了由 (10) 定义的 P 是 $\mathscr{D}$ 上的集函数.

(ii) $P(\Omega)=1, P$ 在 $\mathscr{D}$ 上非负.

(iii) P 满足 (9): 由于有限维乘积概率的性质

$$\begin{aligned}P(A_1\times\cdots\times A_n\times\Omega^{(n)})&=(P_1\times\cdots\times P_n)(A_1\times\cdots\times A_n)\\&=P_1(A_1)\cdots P_n(A_n).\end{aligned}$$

(iv) P 在 $\mathscr{D}$ 上有限可加: 由于 $\mathscr{D}$ 是集代数, 只要证 P 在 $\mathscr{D}$ 上可加.

设 $A,B\in\mathscr{D}$ 且 $A\cap B=\varnothing$, 则存在 n,m 及 $A'\in\mathscr{A}_1\times\cdots\times\mathscr{A}_n, B''\in\mathscr{A}_1\times\cdots\times\mathscr{A}_m$, 使

$$A=A'\times\Omega^{(n)}, B=B''\times\Omega^{(m)}.$$

不妨设 $m\leqslant n$, 这样 B 可改写为

$$B=B'\times\Omega^{(n)},$$

其中 $B'=B''\times\Omega_{m+1}\times\cdots\times\Omega_n\in\mathscr{A}_1\times\cdots\times\mathscr{A}_n$. 由于 $A\cap B=\varnothing$, 故 $A'\cap B'=\varnothing$, 所以

$$P(A+B)=P_1\times\cdots\times P_n(A'+B')$$

$$= (P_1 \times \cdots \times P_n)(A') + (P_1 \times \cdots \times P_n)(B')$$
$$= P(A) + P(B).$$

至此, (一) 的证明已经完成.

(二) 证明 P 在 $\mathscr{D}$ 上在 $\varnothing$ 处连续. 从而由 (一) 及第 1 章 1.4.1 节定理 2 知 P 是 $\mathscr{D}$ 上的概率. 再由测度扩张定理知 P 可以扩张为 $\mathscr{A} = \sigma(\mathscr{D})$ 上的概率.

在证明 (二) 之前, 我们先指出以下事实: 对于 $\mathscr{D}^{(n)}$ 我们可以类似地定义一个集函数 $P^{(n)}$,

$$P^{(n)}(A \times \Omega^{(n+m)}) = (P_{n+1} \times \cdots \times P_{n+m})(A), A \in \mathscr{A}_{n+1} \times \cdots \times \mathscr{A}_{n+m},$$

在 $\mathscr{D}^{(n)}$ 上 $P^{(n)}$ 是有限可加的且满足: 对一切 $A_{n+k} \in \mathscr{A}_{n+k}, k = 1, \cdots, m$ 有

$$P^{(n)}(A_{n+1} \times \cdots \times A_{n+m} \times \Omega^{(n+m)}) = P_{n+1}(A_{n+1}) \cdots P_{n+m}(A_{n+m}).$$

此外, 集函数 P 可以表示成积分形式. 因为对于任意 $(\omega_1, \cdots, \omega_n) \in \Omega_1 \times \cdots \times \Omega_n$, $\mathscr{D}$ 中集合 $A = A' \times \Omega^{(m)} (A' \in \mathscr{A}_1 \times \cdots \times \mathscr{A}_m, m > n)$ 在 $(\omega_1, \cdots, \omega_n)$ 处的截集

$$A(\omega_1, \cdots, \omega_n) = A'(\omega_1, \cdots, \omega_n) \times \Omega^{(m)},$$

其中 $A'(\omega_1, \cdots, \omega_n) \in \mathscr{A}_{n+1} \times \cdots \times \mathscr{A}_m$, 所以

$$P(A) = \int_{\Omega_1 \times \cdots \times \Omega_n} (P_{n+1} \times \cdots \times P_m)(A'(\omega_1, \cdots, \omega_n)) d(P_1 \times \cdots \times P_n)$$

而

$$P^{(n)}(A(\omega_1, \cdots, \omega_n)) = (P_{n+1} \times \cdots \times P_m)(A'(\omega_1, \cdots, \omega_n)).$$

于是得出下面的公式

$$P(A) = \int_{\Omega_1 \times \cdots \times \Omega_n} P^{(n)}(A(\omega_1, \cdots, \omega_n)) d(P_1 \times \cdots \times P_n), \ A \in \mathscr{D}.$$

同样, 对一切 $A \in \mathscr{D}^{(n)}$ 也有

$$P^{(n)}(A) = \int_{\Omega_{n+1} \times \cdots \times \Omega_{n+m}} P^{(n+m)}(A(\omega_{n+1}, \cdots, \omega_{n+m})) d(P_{n+1} \times \cdots \times P_{n+m}).$$

现在我们来证明 (二), 即证明"若 $\{A_n\} \subset \mathscr{D}, A_n \downarrow \varnothing$, 则 $P(A_n) \longrightarrow 0$". 为此, 我们证明它的等价命题:

"若 $\{A_n\} \subset \mathscr{D}, A_n \downarrow$, 且存在一 $\varepsilon > 0$, 使 $P(A_n) \geqslant \varepsilon$ 对一切自然数 n 成立, 则 $\bigcap\limits_{n=1}^{\infty} A_n \neq \varnothing$".

今证在上述条件下存在一个公共元素 $(\bar{\omega}_1,\bar{\omega}_2,\cdots)\in A_n, n=1,2,\cdots$. 首先令

$$B_1^{(j)}=\left\{\omega_1:\omega_1\in\Omega_1, P^{(1)}(A_j(\omega_1))\geqslant\frac{\varepsilon}{2}\right\},$$

显然 $A_j(\omega_1)\supset A_{j+1}(\omega_1)$, 因而 $P^{(1)}(A_j(\omega_1))\geqslant P^{(1)}(A_{j+1}(\omega_1))$, 这就得出 $B_1^{(j)}\downarrow$, 即

$$B_1^{(1)}\supset B_1^{(2)}\supset\cdots.$$

再者, 由于 $A_j\in\mathscr{D}$, 存在 n_j 及 $A_j'\in\mathscr{A}_1\times\cdots\times\mathscr{A}_{n_j}$ 使得 $A_j=A_j'\times\Omega^{(n_j)}$, 因而 $A_j(\omega_1)=A_j'(\omega_1)\times\Omega^{(n_j)}, A_j'(\omega_1)\in\mathscr{A}_2\times\cdots\times\mathscr{A}_{n_j}$, 故

$$P^{(1)}(A_j(\omega_1))=P_2\times\cdots\times P_{n_j}(A_j'(\omega_1))$$

是 $\mathscr{A}_1$-可测函数, 从而 $B_1^{(j)}\in\mathscr{A}_1, j=1,2,\cdots$, 由于

$$\begin{aligned}\varepsilon\leqslant P(A_j)&=\int_{\Omega_1}P^{(1)}(A_j(\omega_1))dP_1\\&=\int_{B_1^{(j)}}P^{(1)}(A_j(\omega_1))dP_1+\int_{(B_1^{(j)})^c}P^{(1)}(A_j(\omega_1))dP_1\\&\leqslant\int_{B_1^{(j)}}dP_1+\int_{\Omega_1}\frac{\varepsilon}{2}dP_1\\&\leqslant P_1(B_1^{(j)})+\frac{\varepsilon}{2},\end{aligned}$$

可知

$$P_1(B_1^{(j)})\geqslant\frac{\varepsilon}{2},\ j=1,2\cdots.$$

因为 P_1 是 $\mathscr{A}_1$ 上的概率, 它在 $\varnothing$ 处连续, 故由此得出 $\bigcap\limits_{j=1}^{\infty}B_1^{(j)}\neq\varnothing$, 这样我们就证明了存在一个 $\bar{\omega}_1\in\bigcap\limits_{j=1}^{\infty}B_1^{(j)}$, 也就是说, 存在一个 $\bar{\omega}_1\in\Omega_1$, 使对一切自然数 $j, P^{(1)}(A_j(\bar{\omega}_1))\geqslant\dfrac{\varepsilon}{2}$.

其次, $A_j(\bar{\omega}_1\in\mathscr{D}^{(1)}, P^{(1)}(A_j(\bar{\omega}_1))\geqslant\dfrac{\varepsilon}{2}, j=1,2,\cdots$ 且 $A_j(\bar{\omega}_1)\downarrow$, 令

$$B_2^{(j)}=\left\{\omega_2:\omega_2\in\Omega_2, P^{(2)}(A_j(\bar{\omega}_1,\omega_2))\geqslant\frac{\varepsilon}{2^2}\right\}.$$

注意到 $A_j(\bar{\omega}_1,\omega_2)=[A_j(\bar{\omega}_1)](\omega_2)$, 重复上面的论证, 又可知存在一 $\bar{\omega}_2\in\Omega_2$, 使 $P^{(2)}(A_j(\bar{\omega}_1,\bar{\omega}_2))\geqslant\dfrac{\varepsilon}{2^2}$, 对一切 $j=1,2,\cdots$ 成立, 且 $A_j(\bar{\omega}_1,\bar{\omega}_2)$ 是 $\mathscr{D}^{(2)}$ 中的不增集

序列. 应用归纳法可以证明存在一 $(\bar{\omega}_1,\bar{\omega}_2,\cdots)\in\Omega$, 使得对一切自然数 n 及 j 有

$$P^{(n)}(A_j(\bar{\omega}_1,\cdots,\bar{\omega}_n))\geqslant\frac{\varepsilon}{2^n}.$$

我们证明这个 $(\bar{\omega}_1,\bar{\omega}_2,\cdots)$ 属于一切 A_j.

事实上, 对于任意 A_j, 由 $A_j\in\mathscr{D}$, 存在 n_j 及 $A_j'\in\mathscr{A}_1\times\cdots\times\mathscr{A}_{n_j}$, 使

$$A_j=A_j'\times\Omega^{(n_j)},$$

由于

$$P^{(n_j)}(A_j(\bar{\omega}_1,\cdots,\bar{\omega}_{n_j}))\geqslant\frac{\varepsilon}{2^{n_j}}>0,$$

故 $A_j(\bar{\omega}_1,\cdots,\bar{\omega}_{n_j})\neq\varnothing$, 而

$$A_j(\bar{\omega}_1,\cdots,\bar{\omega}_{n_j})=\begin{cases}\Omega^{(n_j)}, & 若(\bar{\omega}_1,\cdots,\bar{\omega}_{n_j})\in A_j',\\ \varnothing, & 若(\bar{\omega}_1,\cdots,\bar{\omega}_{n_j})\notin A_j'.\end{cases}$$

由此可见 $(\bar{\omega}_1,\cdots,\bar{\omega}_{n_j})\in A_j'$, 因而对任何的 $(\omega_{n_j+1},\omega_{n_j+2},\cdots)$,

$$(\bar{\omega}_1,\cdots,\bar{\omega}_{n_j},\omega_{n_j+1},\omega_{n_j+2},\cdots)\in A_j,$$

特别

$$(\bar{\omega}_1,\cdots,\bar{\omega}_{n_j},\bar{\omega}_{n_j+1},\cdots)\in A_j.$$

这样我们就证明了 P 在 $\mathscr{D}$ 上在 $\varnothing$ 处连续. 根据前面的分析, 我们就证明了在 $\mathscr{A}$ 上存在一个满足 (9) 的概率.

(三) 最后我们来证明 $\mathscr{A}$ 上满足 (9) 的概率是唯一的. 这是由于 $\mathscr{C}$ 是半集代数,$\sigma(\mathscr{C})=\mathscr{A}$, 满足 (9) 的 $\mathscr{A}$ 上的概率是 $\mathscr{C}$ 上同一概率的扩张, 由测度扩张定理, 它是唯一的. □

在可数无穷乘积概率的构造问题解决之后, 为了构造任意参数集的无穷乘积概率, 我们还需要回过头来再讨论一下可数无穷乘积概率的问题. 在定理 1 中从叙述到证明我们都使用了参数集 $\{1,2,\cdots\}$ 以及它的序. 而实际上可以不考虑这个序.

设 T_c 为任意可数无穷个元素作成的可数集,$(\Omega_t,\mathscr{A}_t,P_t),t\in T_c$ 为可数无穷个概率场. 设 $T_N\subset T_c,T_N=\{t_1,\cdots,t_n\}$ 是 T_c 的任一有限子集, 记

$$\prod_{t\in T_N}\Omega_t=\left\{\left\{\omega_{t_1},\cdots,\omega_{t_n}\right\}:\omega_{t_i}\in\Omega_{t_i},i=1,\cdots,n\right\}$$

它的元素不再是 $(\omega_1,\cdots,\omega_n)$, 而是限定 $\omega_{t_i}\in\Omega_{t_i},i=1,\cdots,n$ 的点集 $\{\omega_{t_1},\cdots,\omega_{t_n}\}$. 它的子集

$$\prod_{t\in T_N}A_t=\left\{\left\{\omega_{t_1},\cdots,\omega_{t_n}\right\}:\omega_{t_i}\in A_{t_i},i=1,\cdots,n\right\}$$

它与我们以前定义的 $A_{t_1}\times\cdots\times A_{t_n}$ 的区别只是不必考虑 $t_1,\cdots,t_n$ 之次序. 则集类

$$\left\{\prod_{t\in T_N}A_t:A_{t_i}\in\mathscr{A}_{t_i},i=1,\cdots,n\right\}$$

是 $\prod\limits_{t\in T_N}\Omega_t$ 中的半集代数, 其上的最小 σ-代数记作

$$\prod_{t\in T_N}\mathscr{A}_t=\sigma\Big(\prod_{t\in T_N}A_t:A_{t_i}\in\mathscr{A}_{t_i},i=1,\cdots,n\Big),$$

与 §4.1 定理 4′ 一样, 令

$$P^{T_N}\Big(\prod_{t\in T_N}A_t\Big)=P_{t_1}\Big(A_{t_1}\Big)\times\cdots\times P_{t_n}\Big(A_{t_n}\Big),$$

则由测度扩张定理 P^{T_N} 可以扩张到 $\prod\limits_{t\in T_N}\mathscr{A}_t$ 上 $\Big($可记 P^{T_N} 为 $\prod\limits_{t\in T_N}P_t\Big)$. 于是就建立了有限维乘积概率空间 $\Big(\prod\limits_{t\in T_N}\Omega_t,\prod\limits_{t\in T_N}\mathscr{A}_t,\prod\limits_{t\in T_N}P_t\Big)$ 可简记为 $(\Omega^{T_N},\mathscr{A}^{T_N},P^{T_N})$.

令

$$\Omega^{T_c}=\prod_{t\in T_c}\Omega_t,$$

$$\mathscr{A}^{T_c}=\prod_{t\in T_c}\mathscr{A}_t=\sigma(B_{T_N}\times\prod_{t\in T_c\setminus T_N}\Omega_t:B_{T_N}\in\prod_{t\in T_N}\mathscr{A}_t,T_N\subset T_c,T_N\text{为有限集}),$$

在可测空间 $(\Omega^{T_c},\mathscr{A}^{T_c})$ 上定义概率, 它是唯一满足

$$P\Big(\prod_{t\in T_N}A_t\times\prod_{t\in T_c\setminus T_N}\Omega_t\Big)=P_{t_1}(A_{t_1})\times\cdots\times P_{t_n}(A_{t_n}),$$

(对任何 $T_N\subset T_c,T_N=\{t_1,\cdots,t_n\},A_{t_i}\in\mathscr{A}_{t_i},i=1,\cdots,n$) 的概率. 记作 $\prod\limits_{t\in T_c}P_t$ 或 P^{T_c}. 于是 $\Big(\prod\limits_{t\in T_c}\Omega_t,\prod\limits_{t\in T_c}\mathscr{A}_t,\prod\limits_{t\in T_c}P_t\Big)$ 是 $(\Omega_t,\mathscr{A}_t,P_t),t\in T_c$ 的乘积概率空间.

这样, 对于任意参数集的情形构造无穷乘积概率的问题就很容易解决了.

设 T 是不可数无穷参数集, $(\Omega_t,\mathscr{A}_t,P_t),t\in T$ 为不可数个概率空间, 自然定义乘积空间为

$$\Omega^T=\prod_{t\in T}\Omega_t=\{\omega_T:\omega_T=\{\omega_t,t\in T\},\omega_t\in\Omega_t,t\in T\}$$

乘积 σ-代数为

$$\mathscr{A}^T = \prod_{t\in T}\mathscr{A}_t = \sigma\Big\{B^{T_N} \times \prod_{t\in T-T_N} \Omega_t : B^{T_N} \in \prod_{t\in T_N} \mathscr{A}_t, T_N \subset T, T_N\text{为有限集}\Big\}.$$

引理 1 设 T 是任一不可数参数集,$(\Omega_t,\mathscr{A}_t), t \in T$ 是一族可测空间, 则在 $\Omega^T = \prod\limits_{t\in T} \Omega_t$ 上的乘积 σ-代数

$$\mathscr{A}^T = \bigcup_{\substack{T_c\subset T\\ T_c\text{可数}}} \Big\{A_{T_c} \times \prod_{t\in T-T_c} \Omega_t : A_{T_c} \in \prod_{t\in T_c} \mathscr{A}_t\Big\}. \tag{11}$$

证 只需证 (11) 式所述集类对可数并及可数交集运算封闭. 设 $B_n = A_{T_{c_n}} \times \prod\limits_{t\in T\setminus T_{c_n}} \Omega_t, A_{T_{c_n}} \in \mathscr{A}^{T_{c_n}}, T_{c_n}, n = 1, 2, \cdots$ 为 T 的可数子集,令 $T_c = \bigcup\limits_{n=1}^{\infty} T_{c_n}$, 则 T_c 仍为 T 的可数子集 (注意: 这里对 T_c 中元无需排序). 则 $B_n = A_{T_{c_n}} \times \prod\limits_{t\in T_c-T_{c_n}} \Omega_t \times \prod\limits_{t\in T-T_c} \Omega_t, n = 1, 2, \cdots$, 而 $A_{T_{c_n}} \times \prod\limits_{t\in T_c-T_{c_n}} \Omega_t \in \mathscr{A}^{T_c}, n = 1, 2, \cdots$. 由于 $\mathscr{A}^{T_c}$ 为 σ-代数, 故它对可数并及可数交集运算封闭, 故有

$$\bigcup_{n=1}^{\infty}\Big(A_{T_{c_n}} \times \prod_{t\in T_c-T_{c_n}} \Omega_t\Big) \in \mathscr{A}^{T_c},$$

$$\bigcap_{n=1}^{\infty}\Big(A_{T_{c_n}} \times \prod_{t\in T_c\setminus T_{c_n}} \Omega_t\Big) \in \mathscr{A}^{T_c}.$$

于是, $\bigcup\limits_{n=1}^{\infty} B_n, \bigcap\limits_{n=1}^{\infty} B_n$ 均为 (11) 式 $\mathscr{A}^T$ 中集. 这就证明了 (11) 式 □

定理 2(无穷乘积概率的存在唯一性定理) 设 $(\Omega_t, \mathscr{A}_t, P_t), t \in T$ 是一族概率空间, 则在乘积可测空间 $\Big(\prod\limits_{t\in T}\Omega_t, \prod\limits_{t\in T}\mathscr{A}_t\Big)$ 上存在唯一的概率 P^T 满足

$$P^T\Big(\prod_{t\in T_N} A_t \times \prod_{t\in T-T_N} \Omega_t\Big) = \prod_{t\in T_N} P_t(A_t), \tag{12}$$

其中 T_N 是 T 中的任一有限集,$A_t \in \mathscr{A}_t, t \in T_N$. 这个概率 P^T 称为 $P_t, t \in T$ 的乘积概率, 记作 $\prod\limits_{t\in T} P_t$, 而 $\Big(\prod\limits_{t\in T}\Omega_t, \prod\limits_{t\in T}\mathscr{A}_t, \prod\limits_{t\in T}P_t\Big)$ 称为 $(\Omega_t, \mathscr{A}_t, P_t), t \in T$ 的乘积概率场, 简记为 $(\Omega^T, \mathscr{A}^T, P^T)$.

证 设 $T_c \subset T, T_c$ 是任意可数集, 根据定理 1 在 $\left(\prod_{t\in T_c} \Omega_t, \prod_{t\in T_c} \mathscr{A}_t\right)$ 上存在可数无穷乘积概率 $\prod_{t\in T_c} P_t$, 将此可数无穷乘积概率空间简记为 $(\Omega^{T_c}, \mathscr{A}^{T_c}, P^{T_c})$, 对于任意的 $A \in \mathscr{A}^T$, 根据 (11) 必存在 T_c 及 $A_{T_c} \in \mathscr{A}^{T_c}$, 使 $A = A_{T_c} \times \prod_{t\in T-T_c} \Omega_t$, 令

$$P^T(A) = P^{T_c}(A_{T_c}). \tag{13}$$

我们先证明由 (13) 定义的 P^T 是 $\mathscr{A}^T$ 上的集函数: 若还有 T_c' 及 $A'_{T_c} \in \mathscr{A}^{T_c'}$, 使 $A = A_{T_c'} \times \prod_{t\in T-T_c'} \Omega_t$, 于是必存在 $A_{T_c\cap T_c'} \in \mathscr{A}^{T_c\cap T_c'}$ 使

$$A_{T_c} = A_{T_c\cap T_c'} \times \prod_{t\in T_c-T_c'} \Omega_t,$$

$$A_{T_c'} = A_{T_c\cap T_c'} \times \prod_{t\in T_c'-T_c} \Omega_t.$$

如果我们证明了

$$P^{T_c}(A_{T_c}) = P^{T_c\cap T_c'}(A_{T_c\cap T_c'}) = P^{T_c'}(A_{T_c'}), \tag{14}$$

则 (13) 确实定义了 $\mathscr{A}^T$ 上的集函数. 为此, 我们将证明一个一般命题, 即: 若 $T_1^* \subset T_2^*, T_1^*, T_2^* \subset T$ 且均为可数集, 则对一切 $A_{T_1^*} \in \mathscr{A}^{T_1^*}$ 有

$$P^{T_1^*}(A_{T_1^*}) = P^{T_2^*}\left(A_{T_1^*} \times \prod_{t\in T_2^*-T_1^*} \Omega_t\right). \tag{15}$$

事实上, 令

$$\mathfrak{M} = \{A_{T_1^*} : A_{T_1^*} \in \mathscr{A}^{T_1^*} \text{且使 (15) 式成立}\}.$$

则 $\mathfrak{M}$ 是包含 π-系

$$\mathscr{C} = \{\prod_{t\in T_N} A_t \times \prod_{t\in T_1^*-T_N} \Omega_t : T_N = \{t_1, \cdots, t_n\} \subset T_1^*, A_{t_i} \in \mathscr{A}_{t_i}, i = 1, \cdots, n\}$$

的 λ-系, 而 $\sigma(\mathscr{C}) = \mathscr{A}^{T_1^*}$, 因而命题获证. (14) 式成立.

再证明由 (13) 式定义的 P^T 是 $\mathscr{A}^T$ 上的概率. 显然 $P^T(\Omega^T) = 1$, 若 $A_n \in \mathscr{A}^T, n = 1, 2, \cdots$ 且两两不交, 则对每个 n, 存在 $T_c^{(n)} \subset T$ 是可数集, 使

$$A_n = A_{T_c^{(n)}} \times \prod_{t\in T-T_c^{(n)}} \Omega_t, A_{T_c^{(n)}} \in \mathscr{A}^{T_c^{(n)}}.$$

令 $T_c = \bigcup_{n=1}^{\infty} T_c^{(n)}$, 则 $T_c \subset T$ 且仍为可数集, 存在 $A_n' \in \mathscr{A}^{T_c}, n = 1, 2, \cdots$ 使

$$A_n = A_n' \times \prod_{t \in T - T_c} \Omega_t, n = 1, 2, \cdots.$$

由于 $A_n, n = 1, 2, \cdots$ 两两不交, 因而 $A_n', n = 1, 2, \cdots$ 两两不交, 且 $\sum_{n=1}^{\infty} A_n' \in \mathscr{A}^{T_c}$. 于是

$$\begin{aligned} P^T\left(\sum_{n=1}^{\infty} A_n\right) &= P^T\left(\left(\sum_{n=1}^{\infty} A_n'\right) \times \prod_{t \in T - T_c} \Omega_t\right) \\ &= P^{T_c}\left(\sum_{n=1}^{\infty} A_n'\right) = \sum_{n=1}^{\infty} P^{T_c}(A_n') \\ &= \sum_{n=1}^{\infty} P^T(A_n). \end{aligned}$$

由 (13) 定义的 P^T 显然满足 (12).

最后证明满足 (12) 的 P^T 是唯一的. 这是由于以可测矩形为底的一切可测柱集的全体构成半集代数 $\mathscr{C}^T$, 而任何满足 (12) 的 $\mathscr{A}^T$ 上的概率在 $\mathscr{C}^T$ 上具有同样的值, 且 $\mathscr{A}^T = \sigma(\mathscr{C}^T)$, 因而由测度扩张定理知满足 (12) 的概率测度 P^T 唯一. □

现在我们来解决本章开始时提出的另一个问题.

例 1 设 $(\Omega, \mathscr{A}, P)$ 是一个概率场, ξ 是它上面的随机变量, 具有分布函数 $F(x)$, 则存在一个概率场, 在其上有独立随机变量序列 $\xi_n, n = 1, 2, \cdots$, 且每一 ξ_n 均与 ξ 有相同的分布函数.

事实上, 我们令

$$\Omega^{\infty} = \Omega_1 \times \Omega_2 \times \cdots = \prod_{n=1}^{\infty} \Omega_n, \text{每一} \Omega_n = \Omega,$$

$$\mathscr{A}^{\infty} = \mathscr{A}_1 \times \mathscr{A}_2 \times \cdots = \prod_{n=1}^{\infty} \mathscr{A}_n, \text{每一} \mathscr{A}_n = \mathscr{A},$$

$$P^{\infty} = P_1 \times P_2 \times \cdots = \prod_{n=1}^{\infty} P_n, \text{每一} P_n = P,$$

则 $(\Omega^{(\infty)}, \mathscr{A}^{(\infty)}, P^{(\infty)})$ 是 $(\Omega_i, \mathscr{A}_i, P_i), i = 1, 2, \cdots$ 的无穷乘积概率空间, 在其上定义

$$\xi_n(\omega_1, \omega_2, \cdots) = \xi(\omega_n), n = 1, 2, \cdots,$$

则 ξ_n 是 $(\Omega^{(\infty)}, \mathscr{A}^{(\infty)}, P^{(\infty)})$ 上的随机变量,$n = 1, 2, \cdots$, 这是因为

$$\{\xi_n < x_n\} = \Omega_1 \times \cdots \times \Omega_{n-1} \times \{\xi(\omega_n) < x_n\} \times \Omega_{n+1} \times \cdots \in \mathscr{A}^{(\infty)}.$$

ξ_n 的分布函数

$$F_{\xi_n}(x_n) = P^{(\infty)}(\xi_n < x_n) = P(\xi < x_n) = F(x_n).$$

即 ξ_n 与 ξ 具有相同的分布函数. 最后证明 $\{\xi_n\}$ 是独立随机变量序列. 事实上, 对任意 k, 任意 $i_1, \cdots, i_k$ 及任意 $x_{i_1}, \cdots, x_{i_k}$,

$$\begin{aligned} &\{\xi_{i_1} < x_{i_1}, \cdots, \xi_{i_k} < x_{i_k}\} \\ = &\prod_{j=1}^{k}\{\xi_{i_j} < x_{i_j}\} = \prod_{j=1}^{k}\{\xi(\omega_{i_j}) < \xi_{i_j}\} \times \prod_{\substack{n \neq i_j \\ j=1,\cdots,k}} \Omega_n. \end{aligned}$$

由此,

$$\begin{aligned} F_{\xi_{i_1},\cdots,\xi_{i_k}}(x_{i_1}, \cdots, x_{i_k}) &= P^{(\infty)}(\xi_{i_1} < x_{i_1}, \cdots, \xi_{i_k} < x_{i_k}) \\ &= \prod_{j=1}^{k} P_{i_j}(\{\omega_{i_j} : \xi(\omega_{i_j}) < x_{i_j}\}) = \prod_{j=1}^{k} P(\xi < x_{i_j}) \\ &= F(x_{i_1}) \cdots F(x_{i_j}) = F_{\xi_{i_1}}(x_{i_1}) \cdots F_{\xi_{i_k}}(x_{i_k}). \end{aligned}$$

所以 $\xi_1, \xi_2, \cdots$ 是独立随机变量序列. □

更一般, 我们有

定理 3 设 $F_n(x_n), n = 1, 2, \cdots$ 是给定的一系列概率分布函数, 那么一定存在一个独立随机变量序列 $\{\xi_n\}, \xi_n$ 的分布函数恰好是 $F_n(x_n)$.

证 设 R 是实数空间, $\mathscr{B}$ 是直线上 Borel 域, 由每一 $F_n(x_n)$ 都决定一个概率空间 $(R, \mathscr{B}, P_n)$, 取它们的乘积概率空间

$$\Omega = R \times R \times \cdots,$$

$$\mathscr{A} = \mathscr{B} \times \mathscr{B} \times \cdots,$$

$$P = P_1 \times P_2 \times \cdots,$$

容易证明

$$\xi_n(x_1, x_2, \cdots) = x_n, \quad n = 1, 2, \cdots$$

是 $(\Omega, \mathscr{A}, P)$ 上的独立随机变量序列, 且对每一自然数 n, ξ_n 的分布函数 $F_{\xi_n}(x_n) = F_n(x_n)$. □

最后, 我们应用无穷乘积概率场建立负二项分布 (参看第 3 章 §3.5 例 2) 的概率场.

例 2 用 $(\omega_k, \bar{\omega}_k), k = 1, 2, \cdots$ 表示一个无穷独立试验序列, 用 ω_k 表示第 k 次试验事件 A_k 发生 (例如在射击试验中表示第 k 次射击击中目标), 且每次试验

事件 A_k 发生的概率相同, 均为 p. 于是我们建立概率场 $(\Omega_k,\mathscr{A}_k,P_k),k=1,2,\cdots$, 其中

$$\Omega_k=\{\omega_k,\bar\omega_k\},$$

$$\mathscr{A}_k=\{\varnothing,\{\omega_k\},\{\bar\omega_k\},\Omega_k\},$$

$$P_k(\{\omega_k\})=p,P_k(\{\bar\omega_n\})=1-p(\text{记作}q).$$

把它们所构成的无穷乘积概率空间 $\left(\prod\limits_{k=1}^{\infty}\Omega_k,\prod\limits_{k=1}^{\infty}\mathscr{A}_k,\prod\limits_{k=1}^{\infty}P_k\right)$ 简记为 $(\Omega,\mathscr{A},P)$, 令

$$\xi_r(\omega_1',\omega_2',\cdots)=\begin{cases}\infty, & \text{若}\omega_1',\omega_2',\cdots\text{中}\omega_i'=\omega_i\text{的个数小于}r,\\ r+k, & \text{若}\omega_{r+k}'=\omega_{r+k},\text{而在}\omega_1',\cdots,\omega_{r+k-1}'\text{中恰有}r-1\text{个}\omega_i\end{cases}$$

若 ξ_r 有限, 则表示第 r 次发生事件 A_i 的试验次数, 而

$$\{\xi_r,=\infty\}=\sum_{m=1}^{r-1}\sum_{1\leqslant i_1<\cdots<i_m}\prod_{s=1}^{m}\{\omega_{i_s}\}\times\prod_{j\neq i_s}\{\bar\omega_j\}$$

是 Ω 中可数个单点集的并. 对每一个单点集,

$$P\left(\prod_{s=1}^{m}\{\omega_{i_s}\}\times\prod_{j\neq i_s}\{\bar\omega_j\}\right)=\lim_{n\to\infty}p^mq^n=0.$$

因而 ξ_r 在一个零测集上取 ∞ 为值, 它 a.e 与一个随机变量相等. 以下计算它的分布列. 由于

$$\{\xi_r=r+k\}=\{(\omega_1',\cdots,\omega_{r+k-1}'):\omega_1',\cdots,\omega_{r+k-1}'$$

$$\text{中恰有}r-1\text{个}\omega_i\}\times\{\omega_{r+k}\}\times\prod_{j=r+k+1}^{\infty}\Omega_j.$$

于是由无穷乘积概率的定义及 §4.1 例 5 知

$$\begin{aligned}P(\xi_r=r+k)&=(P_1\times\cdots\times P_{r+k-1}\times P_{r+k})(A_{r+k-1}^{(r-1)}\times\{\omega_{r+k}\})\\&=(P_1\times\cdots\times P_{r+k-1})(A_{r+k-1}^{(r-1)})\cdot P_{r+k}(\{\omega_{r+k}\})\\&=\begin{pmatrix}r+k-1\\r-1\end{pmatrix}p^{r-1}\cdot q^k\cdot p=\begin{pmatrix}r+k-1\\k\end{pmatrix}p^rq^k.\end{aligned}$$

由于

$$\begin{aligned}\binom{r+k-1}{k} &= \frac{(r+k-1)(r+k-2)\cdots(r+1)r}{k!}\\ &= (-1)^k\frac{(-r)(-r-1)\cdots(-r-(k-1))}{k!}\\ &= (-1)^k\binom{-r}{k},\end{aligned}$$

故

$$P(\xi_r=r+k)=(-1)^k\binom{-r}{k}p^rq^k, k=0,1,2,\cdots$$

□

习题及补充

1. 设 $\xi_1,\xi_2,\cdots$ 是给定的随机变量序列，证明存在独立随机变量序列 $\{\xi_n'\}$, 使 ξ_n' 与 ξ_n 同分布，$n=1,2,\cdots$.

2. 设 T 是任意指标集，对每一 $t\in T, f_t$ 是由 $(\Omega,\mathscr{A})$ 到 $(\Omega_t',\mathscr{A}_t')$ 的可测映射，若令

$$f(\omega)=(f_t(\omega),t\in T),$$

则 f 是 $(\Omega,\mathscr{A})$ 到 $\left(\prod_{t\in T}\Omega_t',\prod_{t\in T}\mathscr{A}_t'\right)$ 的可测映射. (提示：用 λ-系方法).

第 5 章　条件概率与条件数学期望

条件概率与条件数学期望不论在理论上还是在应用上都有着重要的意义. 例如, 马氏过程就用到条件概率的概念, 而某些数理统计中的充分统计量等也是以条件数学期望为重要工具的. 这两个概念虽然与古典概率论中条件概率与条件数学期望有着密切的联系, 但是又有本质的区别. 这两个概念难于用初等的或经典分析的方法做出确切的表达, 需要借助于测度论这个数学工具, 特别是 Radon-Nikodym 定理. 本章的目的就在于给出条件概率与条件数学期望的一般数学定义, 并讨论它们的若干性质.

§5.1　初 等 情 形

在实践中, 常遇到这样的问题: 设 A, B 都是条件组 $\mathscr{C}$ 之下的随机事件, 要求出在事件 B 发生的条件下, 事件 A 的概率, 这个概率我们记作 $P(A|B)$. 下面我们应用频率的方法来研究它, 然后给出这种概率的数学定义.

设 A, B 是条件组 $\mathscr{C}$ 下的随机事件且 $P(B)>0, \mu_A, \mu_B, \mu_{AB}$ 分别是事件 A, B 和 $A\cap B$ 在 $\mathscr{C}$ 的 n 次实现之下的频数, 于是

$$\frac{\mu_{AB}}{n}=\frac{\mu_B}{n}\cdot\frac{\mu_{AB}}{\mu_B}.$$

由于 $P(B)>0$, 故当 n 充分大时, μ_B 一般来说也可以充分大. 而 μ_{AB} 可以看成是在 B 发生 μ_B 次(即条件组 $\mathscr{C}$ 及“B 发生”实现 μ_B 次) 之下事件 A 的频数. 因此, $\frac{\mu_{AB}}{\mu_B}$ 则成为与 $P(A|B)$ 相应的频率. 所以, 我们有

$$P(A\cap B)=F(B)P(A|B),$$

亦即

$$P(A|B)=\frac{P(A\cap B)}{P(B)},$$

其中 $P(A\cap B), P(B)$ 分别是事件 $A\cap B, B$ 在条件组 $\mathscr{C}$ 之下的概率. 我们有

定义 1　设 $(\Omega, \mathscr{A}, P)$ 是一概率场,$B\in\mathscr{A}$ 且 $P(B)>0$, 对于 $\mathscr{A}$ 中的任一事件 A, 令

$$P(A|B)=\frac{P(A\cap B)}{P(B)}, \tag{1}$$

则 $P(\cdot|B)$ 是定义在 $\mathscr{A}$ 上的集函数, 称之为**在给定事件 B 之下的条件概率**. 对于给定的 $A, P(A|B)$ 称为**事件 A 在给定事件 B 之下的条件概率**.

记 $P_B(A) \widehat{=} P(A|B)$, 则易证 $(\Omega, \mathscr{A}, P_B)$ 是一个概率场 (参看第 1 章 §1.4 习题 8).

性质 1(乘法公式)　若 $A, B \in \mathscr{A}, P(A) > 0, P(B) > 0$, 则

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B). \tag{2}$$

一般, 若 $P(A_1 \cap \cdots \cap A_{n-1}) > 0$, 则

$$P(A_1 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1}). \tag{3}$$

性质 2(全概率公式)　若 $A \in \mathscr{A}, \{B_k\}$ 为 $\mathscr{A}$ 中有限个或可数个两两不相交且具有正概率的事件, $A \subset \sum\limits_k B_k$, 则

$$P(A) = \sum_k P(B_k)P(A|B_k). \tag{4}$$

性质 3(Bayes 公式)　若 $A \in \mathscr{A}, P(A) > 0$, 且 $\{B_k\}$ 是 $\mathscr{A}$ 中有限个或可数个两两不相交的正概率事件, $A \subset \sum\limits_k B_k$, 则对每一 B_k,

$$P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum\limits_n P(B_n)P(A|B_n)}. \tag{5}$$

以上性质的证明较易, 希读者完成.

既然 $(\Omega, \mathscr{A}, P_B)$ 是一个概率场, 我们就可以考虑随机变量 ξ 在此概率场上的积分, 这就导出

定义 2　设 ξ 是概率场 $(\Omega, \mathscr{A}, P_B)$ 上的随机变量且 $\int_\Omega \xi dP_B$ 存在, 则称它为 ξ 在给定事件 B 之下的条件数学期望, 记作 $E(\xi|B)$.

性质 4　若 $E\xi$ 存在 (不一定有限), $P(B) > 0$, 则 $E(\xi|B)$ 存在, 且

$$E(\xi|B) = \frac{1}{P(B)} \int_B \xi dP. \tag{6}$$

证　当 $\xi = \chi_A, A \in \mathscr{A}$ 时, 显然

$$\begin{aligned} E(\xi|B) &= \int_\Omega \chi_A dP_B = P_B(A) \\ &= \frac{P(A \cap B)}{P(B)} = \frac{1}{P(B)} \int_B \chi_A dP. \end{aligned}$$

于是当 ξ 是非负简单函数时 (6) 成立, 因而当 ξ 为非负随机变量时由单调收敛定理知 (6) 成立. 最后对 $E\xi$ 存在的情形分别考察 ξ^+,ξ^- 之后即知性质 4 成立. □

从性质 4 我们得到两点启示:

1° ξ 在给定事件 B 下的条件数学期望的实际意义是 ξ 在集 B 上的平均值.

2° 事件 A 在给定事件 B 下的条件概率可以看作条件期望的特殊情形 (我们以下简称条件数学期望为条件期望), 即

$$E(\chi_A|B) = P(A|B), A \in \mathscr{A}.$$

其中 1° 对于推广条件期望的概念很有用, 而 2° 使得我们致力于条件期望性质的研究, 而把条件概率的性质作为它的特殊情形顺便推导出来.

但是实际上, 只是孤立地去讨论单个事件下的条件期望还是不够的, 而必须把条件期望了解为 ω 的一个函数. 例如我们同时考虑 $E(\xi|B)$ 和 $E(\xi|B^c)$ 时, 我们不把它们看作孤立的两个数, 而了解为一个函数, 在 $\omega \in B$ 时, 取 ξ 在 B 上的平均值即 $E(\xi|B)$, 而在 $\omega \in B^c$ 时, 取 ξ 在 B^c 上的平均值 $E(\xi|B^c)$. 更一般些, 当 $\{B_n\}$ 是 Ω 的可数分割且 $P(B_n) > 0, n = 1, 2, \cdots$ 时, 我们认定函数

$$\sum_n E(\xi|B_n)\chi_{B_n}(\omega)$$

为一定意义下的条件期望. 为了能够了解它的意义, 我们下面举例说明.

例 1 求在单位时间内某一平面区域 G 上接受来自宇宙质点的平均能量.

我们自然设单位时间落入 G 的质点数为 μ, 第 k 个质点所具有的能量是 ξ_k, ξ_k 及 μ 都是随机变量, 且 μ 以非负整数为值. 于是单位时间内 G 所接受的能量

$$\xi = \xi_1 + \cdots + \xi_\mu,$$

而所需求的答案是 $E\xi$.

当我们看到 ξ 是随机变量的和, 很可能立刻想到是否可以直接应用第 3 章的方法, 将 $E\xi$ 表成 $E\xi_k$ 的和. 进一步当注意到这个和的项数也是随机变量时, 我们就又会感到这样做是困难的. 但是我们可以将它化为求条件期望的问题, 即

$$\begin{aligned} E\xi &= \int_\Omega \xi dP = \sum_{n=1}^{\infty} \int_{\{\mu=n\}} \xi dP \\ &= \sum_n E(\xi|\{\mu = n\})P(\{\mu = n\}) \\ &= \int \sum_n E(\xi|\{\mu = n\})\chi_{\{\mu=n\}} dP. \end{aligned}$$

因此, 求 ξ 的数学期望就化作求条件期望

$$\sum_n E(\xi|\{\mu=n\})\chi_{\{\mu=n\}}$$

的数学期望的问题. (注意到当 $P(\{\mu=n\})=0$ 时 $E(\xi|\{\mu=n\})$ 无意义, 故上述函数可能在一个零概率集上没有定义, 但它不影响积分的结果.)

这个问题的进一步讨论留作习题.

定义 3　设 $(\Omega,\mathscr{A},P)$ 为一概率场, $\{B_n\}\subset\mathscr{A}$ 为 Ω 的可数分割, 令 $\mathscr{C}=\sigma(B_n,n=1,2,\cdots),E\xi$ 存在, 按等价意义 (即可以在 $\mathscr{C}$ 的零概率集上任意取值) 定义的下列 $\mathscr{C}$ 可测函数

$$E(\xi|\mathscr{C})=\sum_{n=1}^{\infty}E(\xi|B_n)\chi_{B_n} \tag{7}$$

叫做 ξ 在给定 σ-代数 $\mathscr{C}$ 之下关于 P 的条件数学期望.

从定义 3 看到 $E(\xi|\mathscr{C})$ 在 $\mathscr{C}$ 的每个原子 B_n 上取 ξ 在此原子上的平均值, 因此 $E(\xi|\mathscr{C})$ 在 $\mathscr{C}$ 的原子上把 ξ 平滑化了.

这里, 我们说"给定 σ-代数 $\mathscr{C}$ 下的条件期望"是因为 $E(\xi|\mathscr{C})$ 不仅可以定出给定事件 B_n 之下 ξ 的条件期望，而且可以定出给定 $\mathscr{C}$ 可测的任意集 $B(P(B)>0)$ 时 ξ 的条件期望 $E(\xi|B)$. 事实上，若 $B\in\mathscr{C}$，则 B 必然是 $\{B_n\}$ 中某些集的不交并，例如说 $B=\sum' B_n$, 于是由性质 4 知

$$\begin{aligned}P(B)E(\xi|B)&=\int_B\xi dP=\sum{}'\int_{B_n}\xi dP\\&=\sum{}'P(B_n)E(\xi|B_n)=\int_B\sum_n E(\xi|B_n)\chi_{B_n}dP\\&=\int_B E(\xi|\mathscr{C})dP,\end{aligned} \tag{8}$$

即

$$E(\xi|B)=\frac{1}{P(B)}\int_B E(\xi|\mathscr{C})dP.$$

不仅如此, (8) 式还可以做为条件期望 $E(\xi|\mathscr{C})$ 的描述性定义, 而定义 3 称为条件期望的构造性定义.

定义 3′　设 ξ 为 $(\Omega,\mathscr{A},P)$ 上的随机变量, $E\xi$ 存在, $\{B_n\}\subset\mathscr{A}$ 为 Ω 的可数分割, $\mathscr{C}=\sigma(B_n,n\geqslant1)$, 称满足下式的 $\mathscr{C}$ 可测函数 $E(\xi|\mathscr{C})$ 为 ξ 在 σ-代数 $\mathscr{C}$ 下关于 P 的条件期望:

$$\int_B\xi dP=\int_B E(\xi|\mathscr{C})dP,\quad B\in\mathscr{C}. \tag{9}$$

由于不定积分 $\varphi(B)=\int_B \xi dP, B\in\mathscr{C}$ 是 $\mathscr{C}$ 上 σ-可加, P-连续的集函数, 因而根据推广的 Radon-Nikodym定理知满足 (9) 的 $\mathscr{C}$ 可测函数存在且在等价意义上是唯一的. 即 $E(\xi|\mathscr{C})=\dfrac{d\varphi}{dP}$.

(8) 式表明, 定义 3 定义的 $E(\xi|\mathscr{C})$ 与 (9) 式确定的 $E(\xi|\mathscr{C})$ 是同一的 (在等价意义下).

对于给定 σ-代数下事件 A 的条件概率同样可以给出上述构造性定义和描述性定义.

定义 3′ 对于进一步推广条件期望提供了途径. 因为在一般情形我们不能要求 $\mathscr{A}$ 的子 σ-代数 $\mathscr{C}$ 是由 Ω 的可数分割生成的最小 σ-代数. 例如 $\mathscr{B}^{(1)}$ 在某一连续型随机变量 η 之下的逆象所作成的最小 σ-代数 $\sigma(\eta)$ 就不是 Ω 的可数分割, 在这种情形下, 我们欲求 $E(\xi|\sigma(\eta))$, 定义 3 就不能使用了, 但定义 3′ 却给了我们解决问题的途径, 在下一节我们就按这种办法给出条件期望的一般定义.

习题及补充

1. 设有 N 个产品, 其中有 M 个废品, 求在第一次抽得合格品的条件下第二次抽得废品的概率.

2. 设某工厂的一种产品是由三台机器生产的, 根据单独测定的结果, 在正常情况下, 三台机器生产的废品率各为 0.01,0.02 和 0.04, 而在同一时间内生产产品数的比是 8:5:2, 试求该厂生产此种产品的废品率.

3. 某工厂产品的合格率是 0.96, 为了保证质量必须进行破坏性的检查. 提出一种检查简化方案, 经试验知, 一个合格品经该方案检查而获出厂的概率是 0.98, 而一个废品经此方案检查而获出厂的概率是 0.05, 问用这种简化方案作为出厂产品检查方案, 获得出厂许可的产品是合格品的概率以及未获出厂许可的产品是废品的概率.

4. 若 ξ 是按具参数 λ 的 Poisson 分布的随机变量, η 在给定事件 $\{\xi=n\}$ 之下的条件分布为具有参数 n,p 的二项分布, 即

$$P(\eta=m|\{\xi=n\})=\begin{pmatrix} n \\ m \end{pmatrix} p^m(1-p)^{n-m}, m=0,1,\cdots,n.$$

证明: η 是参数为 λp 的 Poisson 分布的随机变量.

5. 设 $\xi_1,\xi_2,\cdots,\mu$ 是 $(\Omega,\mathscr{A},P)$ 上的随机变量, μ 只取非负整数值. $\xi=\xi_1+\cdots+\xi_\mu$, 又设 μ 与 $\xi_1,\xi_2,\cdots$ 独立, 试证

$$E(\xi|\{\mu=n\})=\sum_{k=1}^{n} E\xi_k.$$

进一步设级数 $\sum_{k=1}^{\infty} E|\xi_k|P(\mu \geqslant k) < \infty$ 或 ξ_k 非负, $k=1,2,\cdots$ 则 $E\xi$ 存在且

$$E\xi = \sum_{k=1}^{\infty} E\xi_k P(\mu \geqslant k).$$

若对一切 $k \geqslant 1, E\xi_k = a$, 则 $E\xi = a \cdot E\mu$.

6. 用户在单位时间内对电话局要求通话的总时间的平均值为该电话局的话务量. 设单位时间内用户向电话局呼唤的次数 μ 是按具有参数 λ 的 Poisson 律分布的, 而每一用户通话时间 ξ_k 是按负指数分布的, 即

$$P(\xi \geqslant x) = \begin{cases} 1, & x \leqslant 0, \\ e^{-\beta x}, & x > 0, \end{cases} \quad (\beta > 0),$$

则该电话局的话务量是 $\frac{\lambda}{\beta}$.

§5.2　给定 σ- 代数下条件期望与条件概率的定义和性质

5.2.1　定义及简单性质

在前一节分析的基础上, 我们给出下面的

定义 1　设 $(\Omega, \mathscr{A}, P)$ 是一概率场,$\mathscr{C}$ 是 $\mathscr{A}$ 的子 σ-代数, ξ 是 $(\Omega, \mathscr{A}, P)$ 上的随机变量且 $E\xi$ 存在 (不一定有限). ξ 在 $\mathscr{C}$ 之下 (关于 P) 的条件数学期望, 记作 $E(\xi|\mathscr{C})$, 指的是满足下述条件的 Ω 上的一个 $\mathscr{C}$ 可测函数的等价类中的任何一个.

$$\int_B E(\xi|\mathscr{C})dP = \int_B \xi dP, \text{对一切} B \in \mathscr{C}. \tag{1}$$

由于

$$\varphi(B) = \int_B \xi dP, \quad B \in \mathscr{C}$$

是$\mathscr{C}$ 上 σ-可加、P-连续的集函数, 因而由推广的 Radon-Nikodym 定理知 $E(\xi|\mathscr{C}) = \frac{d\varphi}{dP}$ 是存在的且在 $\mathscr{C}$-可测函数等价的意义上是唯一确定的. (注意不难将 Radon-Nikodym 定理推广至 φ 是复的集函数的情形. 这里 ξ 可以是复随机变量.)

关于定义 1, 我们做以下几点重要的说明:

1° 在定义 1 中要求 ξ 是随机变量且 $E\xi$ 存在, 因而不定积分 $\varphi(B) = \int_B \xi dP$ 在 $\mathscr{A}$ 上是 σ-有限的, 但未必在其子 σ-代数 $\mathscr{C}$ 上 σ-有限. 例如 $\mathscr{C} = \{\phi, \Omega\}, \varphi(\Omega) =$

$E\xi=\infty$. 由此可见, 尽管 ξ 是随机变量,$E(\xi|\mathscr{C})$ 也未必是随机变量. (因为当 $E\xi=\infty$ 时 $E(\xi|\mathscr{C})$ 未必 a.e. 有限.)

2° 考察定义 1, 若 ξ 是任意积分存在的可测函数, 满足 (1) 式的 $\mathscr{C}$-可测函数仍然存在且 a.e. 唯一确定. 因此可以把条件期望的定义扩展到任意积分存在的可测函数上, 这时将 $E\xi$ 理解为 $\int\xi dP$. 事实上, 今后要用到这一点, 例如有时需要考虑 $E(\xi|\mathscr{C})$ 在某一 σ-代数 $\mathscr{C}'$ 下的条件期望 $E[E(\xi|\mathscr{C})|\mathscr{C}']$, 这时 $E(\xi|\mathscr{C})$ 就不一定是随机变量了. 但是我们仍然采取概率论中随机变量的语言及记号. 本节及下一节的所有定理都可以推广到积分存在的可测函数上, 请读者注意考察.

3° 在这一节以及以后各节谈及条件期望与条件概率时均指关于一个给定的概率 P 而言的. 为了叙述简单, 我们一般省去"关于 P"的叙述.

4° 由于只有对 $E\xi$ 存在的 $\xi,E(\xi|\mathscr{C})$ 才有意义, 因此今后假定出现在条件期望符号下的可测函数的数学期望 (即积分) 永远存在.

5° 在谈到 $E(\xi|\mathscr{C})$ 的 a.e. 性质时, 例外集均指 $\mathscr{C}$ 可测集, 即存在 $\mathscr{C}$-可测集 $N,P(N)=0$, 使当 $\omega\in N^c$ 时该性质成立. 若将 P 在 $\mathscr{C}$ 上的限制记作 $P_{\mathscr{C}}$, 则可记作, $P_{\mathscr{C}}$-a.e. 成立.

6° 记 $\mathscr{E}=\{\xi: E\xi\text{存在},\xi\text{为}(\Omega,\mathscr{A},P)$ 上的可测函数, 把等价的可测函数看作同一元$\}$, $\mathscr{G}=\{f: f$ 是 $(\Omega,\mathscr{C},P_{\mathscr{C}})$ 中的可测函数的等价类$\}$, 则 $E(\cdot|\mathscr{C})$ 是 $\mathscr{E}$ 到 $\mathscr{G}$ 的映射. 我们把记号 $E(\cdot|\mathscr{C})$ 叫做给定 σ 代数 $\mathscr{C}$ 的条件期望. 另外也可以把 $E(\cdot|\mathscr{C})(\cdot)$ 看作是定义于 $\Omega\times\mathscr{E}$ 上取值于 $\bar{R}$ 中的函数, 在 (ω,ξ) 处之值就是 $E(\xi|\mathscr{C})(\omega)$.(但对一个给定的 $\xi,E(\xi|\mathscr{C})(\cdot)$ 是 $P_{\mathscr{C}}$-a.e. 唯一确定的.)

条件期望 $E(\xi|\mathscr{C})$ 有下列简单性质:

性质 1 若 $E\xi$ 存在, 则 $E[E(\xi|\mathscr{C})]$ 存在且

$$E[E(\xi|\mathscr{C})]=E\xi. \tag{2}$$

证 在 (1) 中令 $B=\Omega$ 即得.

性质 2 若 $\mathscr{C}=\mathscr{A}$ 或 ξ 是 $\mathscr{C}$ 可测函数时, 则 $E(\xi|\mathscr{C})=\xi,P_{\mathscr{C}}$-a.e..

证 在两种假设下, 任何一个与 $\xi P_{\mathscr{C}}$-a.e. 相等的可测函数都满足 (1), 故 $E(\xi|\mathscr{C})=\xi,P_{\mathscr{C}}$-a.e..

性质 3 若 ξ 是复可测函数,$\xi=\xi_1+i\xi_2,\xi_1,\xi_2$ 分别是 ξ 的实部和虚部, 则

$$\begin{aligned}E(\xi|\mathscr{C})&=E(\xi_1|\mathscr{C})+iE(\xi_2|\mathscr{C})\\&=E(\xi_1^+|\mathscr{C})-E(\xi_1^-|\mathscr{C})+i[E(\xi_2^+|\mathscr{C})-E(\xi_2^-|\mathscr{C})].\end{aligned}$$

当把 ξ 局限于 $\mathscr{A}$ 可测集的示性函数时，就得到了给定 σ 代数 $\mathscr{C}$ 下条件概率的概念，即

定义 2　设 $(\Omega, \mathscr{A}, P)$ 是一概率场,$\mathscr{C}$ 是 $\mathscr{A}$ 的子 σ-代数, $A \in \mathscr{A}$, 则称 $E(\chi_A|\mathscr{C})$ 为事件 A 在 σ-代数 $\mathscr{C}$ 下 (关于 P) 的条件概率. 记作 $P(A|\mathscr{C})$.

由性质 1, 2 立刻得出

性质 1′　$E(P(A|\mathscr{C})) = P(A)$.

性质 2′　若 $A \in \mathscr{C}, P(A|\mathscr{C}) = \chi_A, P_{\mathscr{C}}$-a.e..

为了应用, 我们进一步指出: $P(A|\mathscr{C})$ 是满足下述关系的 $P_{\mathscr{C}}$-a.e. 确定的 $\mathscr{C}$-可测函数

$$P(A \cap B) = \int_B P(A|\mathscr{C}) dP_{\mathscr{C}}, B \in \mathscr{C}. \tag{3}$$

条件期望有很多重要性质. 这些性质中有一部分与数学期望作为积分的性质是对应的. 而另一部分所谓平滑性质则与数学期望取“平均值”的意义有一定联系. 这些也是条件数学期望这一名称的由来. 以下分别考虑这两类性质.

5.2.2　给定 σ-代数下条件期望的期望性质

我们将看到, 条件期望几乎具有数学期望的所有性质. 首先是线性及单调性.

定理 1　设 a, b 是常数, 则

I) 若 $\xi = a, P$-a.e., 则 $E(\xi|\mathscr{C}) = a, P_{\mathscr{C}}$-a.e.

II) 若 $aE\xi + bE\eta$ 存在, 则

$$E(a\xi + b\eta|\mathscr{C}) = aE(\xi|\mathscr{C}) + bE(\eta|\mathscr{C}), \quad P_{\mathscr{C}}\text{-a.e.}$$

III) 若 $\xi \leqslant \eta, P$-a.e., 则 $E(\xi|\mathscr{C}) \leqslant E(\eta|\mathscr{C}), P_{\mathscr{C}}$-a.e. 若 ξ 是复可测函数, 则 $|E(\xi|\mathscr{C})| \leqslant E(|\xi||\mathscr{C}), P_{\mathscr{C}}$-a.e.

证　定义 1 中的 (1) 式被 $E(\xi|\mathscr{C}) = a$ 满足, 故 I) 成立. 其次, 由定义 1 及积分线性性知,$aE\xi + bE\eta$ 存在可推得 $E(a\xi + b\eta)$ 存在, 因而对任何的 $B \in \mathscr{C}$

$$\begin{aligned}
\int_B E(a\xi + b\eta|\mathscr{C}) dP &= \int_B (a\xi + b\eta) dP \\
&= a\int_B \xi dP + b\int_B \eta dP \\
&= a\int_B E(\xi|\mathscr{C}) dP + b\int_B E(\eta|\mathscr{C}) dP \\
&= \int_B [aE(\xi|\mathscr{C}) + bE(\eta|\mathscr{C})] dP.
\end{aligned}$$

(注意上式中各项的存在性都是由 $E\xi, E\eta, aE\xi + bE\eta$ 的存在所保证的.) 这就是说 $E(a\xi + b\eta|\mathscr{C})$ 及 $aE(\xi|\mathscr{C}) + bE(\eta|\mathscr{C})$ 在 $\mathscr{C}$ 上的不定积分定义了 $\mathscr{C}$ 上的同一个 σ-可加且 $P_{\mathscr{C}}$-连续的集函数, 因此由推广的 Radon-Nikodym 定理知 II) 成立.

再次, 由于积分性质及定义 1 知对一切 $B\in\mathscr{C}$,

$$\int_B E(\xi|\mathscr{C})dP=\int_B \xi dP\leqslant\int_B \eta dP=\int_B E(\eta|\mathscr{C})dP.$$

因而 $E(\xi|\mathscr{C})\leqslant E(\eta|\mathscr{C}), P_{\mathscr{C}}$-a.e.. (注意这里用到了第 3 章 3.2.2 定理 2,2)a) 的全部结果.) 因而 III) 的第一部分获证.

最后证明 III) 的第二部分. 先设 ξ 是有界随机变量,

$$E(\xi|\mathscr{C})(\omega)=r(\omega)e^{i\theta(\omega)},$$

其中当 $r(\omega)=0$ 时取 $e^{i\theta(\omega)}=1$. 并设 $\xi=\eta+i\zeta,\eta,\zeta$ 是实随机变量. 则由 II) 知

$$r=|E(\xi|\mathscr{C})|=\sqrt{[E(\eta|\mathscr{C})]^2+[E(\zeta|\mathscr{C})]^2}$$

是 $\mathscr{C}$-可测函数, 于是由

$$e^{-i\theta(\omega)}=\begin{cases}\dfrac{r(\omega)}{E(\xi|\mathscr{C})(\omega)}, & r(\omega)\neq 0,\\ 1, & r(\omega)=0.\end{cases}$$

知 $e^{-i\theta}$ 也是 $\mathscr{C}$-可测函数. 由后面将要讲到的条件期望的平滑性质 (§2.3 定理 7) 知

$$|E(\xi|\mathscr{C})|=E(\xi|\mathscr{C})\mathrm{e}^{-i\theta}=E(\xi e^{-i\theta}|\mathscr{C})\geqslant 0.$$

故对任何 $B\in\mathscr{C}$,

$$\begin{aligned}0&\leqslant\int_B|E(\xi|\mathscr{C})|dP=\int_B\xi e^{-i\theta}dP\\&=\left|\int_B\xi e^{-i\theta}dP\right|\leqslant\int_B\Big|\xi\Big|dP\\&=\int_B E(|\xi||\mathscr{C})dP.\end{aligned}$$

因此, III) 的第二部分对 ξ 是有界随机变量时成立.

若 ξ 是 $E\xi$ 存在的有限复函数, 设 $\xi=\xi_1+i\xi_2$, 则 $E\xi_i^+,E\xi_i^-$ 至少有一有限, $i=1,2$. 不妨假定 $E\xi_i^-$ 有限,$i=1,2$. 由定理 2 将要证明的对条件期望的单调收敛定理知

$$\lim_{n\to\infty}E(\xi_i^{\pm}\chi_{\{|\xi|\leqslant n\}}|\mathscr{C})=E(\xi_i^{\pm}|\mathscr{C}),P_{\mathscr{C}}\text{-a.e.},i=1,2.$$

且 $E(\xi_i^-|\mathscr{C}),P_{\mathscr{C}}$-a.e. 有限, $i=1,2$.

由前面的证明知, 对任何正整数 n

$$|E(\xi\chi_{\{|\xi|\leqslant n\}}|\mathscr{C})|\leqslant E(|\xi|\chi_{\{|\xi|\leqslant n\}}|\mathscr{C})\leqslant E(|\xi||\mathscr{C}),P_{\mathscr{C}}\text{-a.e.},$$

而

$$\begin{aligned}
&\lim_{n\to\infty} E(\xi\chi_{\{|\xi|\leqslant n\}}|\mathscr{C}) = \lim_{n\to\infty}\{E(\xi_1^+\chi_{\{|\xi|\leqslant n\}}|\mathscr{C})\\
&\quad - E(\xi_1^-\chi_{\{|\xi|\leqslant n\}}|\mathscr{C}) + i[E(\xi_2^+\chi_{\{|\xi|\leqslant n\}}|\mathscr{C}) - E(\xi_2^-\chi_{\{|\xi|\leqslant n\}}|\mathscr{C})]\}\\
&= E(\xi_1^+|\mathscr{C}) - E(\xi_1^-|\mathscr{C}) + i[E(\xi_2^+|\mathscr{C}) - E(\xi_2^-|\mathscr{C})]\\
&= E(\xi|\mathscr{C}).
\end{aligned}$$

因而

$$|E(\xi|\mathscr{C})| \leqslant E(|\xi||\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}$$

对 ξ 取非有限值的情形, 可通过令 $\xi = \xi\chi_{\{|\xi|<\infty\}} + \xi\chi_{\{|\xi|=\infty\}}$ 而证明, 不再赘述. (前一步证明方法称为截割法.) □

由 II) 知 $E(\cdot|\mathscr{C})$ 是由 $\mathscr{E}$ 到 $\mathscr{G}$(§2.1 所述) 中的线性映射.

在定理 1 中令 $\xi = \chi_A, A \in \mathscr{A}$, 则立得

定理 1′　在 $\mathscr{A}$ 的子 σ-代数 $\mathscr{C}$ 之下的条件概率具有下述性质:

I′)　$P(\Omega|\mathscr{C}) = 1, P_{\mathscr{C}}$-a.e.

$P(\varnothing|\mathscr{C}) = 0, P_{\mathscr{C}}$-a.e.

II′) 若 $A_k \in \mathscr{A}, k = 1, 2, \cdots$, 且两两不交, 则

$$P\left(\sum_{k=1}^{\infty} A_k|\mathscr{C}\right) = \sum_{k=1}^{\infty} P(A_k|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}$$

III′) 对任何 $A \in \mathscr{A}$,

$$0 \leqslant P(A|\mathscr{C}) \leqslant 1, P_{\mathscr{C}}\text{-a.e.}$$

注　II′) 的证明用到了下面定理 2 中的单调收敛性.

应该指出: II′)III′) 所出现的条件概率是对给定的 A_k, A 来说的函数, 因此 $P_{\mathscr{C}}$-a.e. 成立的确切含意是对给定的 A_k(或 A), 有一与 A_k(或 A) 有关的 $\mathscr{C}$ 可测的零概率集 N, 使得 II′)III′) 对一切 $\omega \in N^c$ 成立. 这种零概率集一般来说是随 A_k(或 A) 而变动的. 确切理解定理 1′ 的含意, 对于理解 §5.5 中提出的正则条件概率的概念是十分必要的.

其次我们会看到条件期望保持收敛的性质 (这些性质与积分收敛定理相当), 即

定理 2

IV)(单调收敛性). 若 $0 \leqslant \xi_n \uparrow \xi, P$-a.e., 则

$$0 \leqslant E(\xi_n|\mathscr{C}) \uparrow E(\xi|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}$$

V)(Fatou-Lebesgue 收敛性). 设 η,ζ 可积, 若对一切 $n\geqslant 1,\eta\leqslant\xi_n, P$-a.e., 则

$$E(\varliminf_{n\to\infty}\xi_n|\mathscr{C})\leqslant\varliminf_{n\to\infty}E(\xi_n|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}$$

若对一切 $n\geqslant 1,\xi_n\leqslant\zeta, P$-a.e., 则

$$\varlimsup_{n\to\infty}E(\xi_n|\mathscr{C})\leqslant E(\varlimsup_{n\to\infty}\xi_n|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}$$

VI)(控制收敛性). 设 η,ζ 可积, 若 $\eta\leqslant\xi_n\uparrow\xi, P$-a.e., 或对一切 $n\geqslant 1,\eta\leqslant\xi_n\leqslant\zeta$ 且 $\xi_n\to\xi, P$-a.e. 则

$$E(\xi_n|\mathscr{C})\longrightarrow E(\xi|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}$$

证 i) 由 III) 知, 若 $0\leqslant\xi_n\uparrow\xi, P$-a.e., 则序列 $\{E(\xi_n|\mathscr{C})\}$ 是 $P_{\mathscr{C}}$-a.e. 不降且非负的, 因此有一 $\mathscr{C}$ 可测函数 ξ' 使

$$0\leqslant E(\xi_n|\mathscr{C})\uparrow\xi', P_{\mathscr{C}}\text{-a.e.}$$

于是由积分的单调收敛定理知, 对任意的 $B\in\mathscr{C}$, 有

$$\lim_{n\to\infty}\int_B E(\xi_n|\mathscr{C})dP=\int_B\xi' dP.$$

另一方面, 由定义 1 及单调收敛定理知对任一 $B\in\mathscr{C}$,

$$\lim_{n\to\infty}\int_B E(\xi_n|\mathscr{C})dP=\lim_{n\to\infty}\int_B\xi_n dP=\int_B\xi dP=\int_B E(\xi|\mathscr{C})dP.$$

因此对任一 $B\in\mathscr{C}$, 都有

$$\int_B\xi' dP=\int_B E(\xi|\mathscr{C})dP.$$

由于 ξ' 与 $E(\xi|\mathscr{C})$ 都是 $\mathscr{C}$ 可测函数, 故 $\xi'=E(\xi|\mathscr{C}), P_{\mathscr{C}}$-a.e., 因而

$$0\leqslant E(\xi_n|\mathscr{C})\uparrow E(\xi|\mathscr{C}), P_{\mathscr{C}}\text{-a.e..}$$

ii) V) 的证明与 Fatou-Lebesgue 定理 (第 3 章 §3.1 定理 1) 的证明是一样的, 只需将应用单调收敛定理之处相应地换成 IV 即可. 今简述如下. 首先注意, 因为 η 可积, 故 $E(\eta|\mathscr{C}), P_{\mathscr{C}}$-a.e. 有限, 令

$$\eta_n=\inf_{k\geqslant n}(\xi_k-\eta), n\geqslant 1.$$

于是

$$0 \leqslant \eta_n \uparrow \sup_n \inf_{k \geqslant n} (\xi_k - \eta) = \varliminf_{n\to\infty} (\xi_n - \eta), P\text{-a.e.}$$

故由 III 及 IV 知

$$\begin{aligned} E(\varliminf_{n\to\infty} \xi_n - \eta|\mathscr{C}) &= \lim_{n\to\infty} E(\eta_n|\mathscr{C}) \\ &\leqslant \sup_n \inf_{k\geqslant n} E(\xi_k - \eta|\mathscr{C}) \\ &= \varliminf_{n\to\infty} E(\xi_k - \eta|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}. \end{aligned}$$

在上式两端加上 $E(\eta|\mathscr{C})$, 即得 V 的第一部分. 同法可得第二部分.

iii) VI 是 IV 和 V 的直接推论. □

条件数学期望具有和数学期望相应的一系列不等式, 即

定理 3

VII)(C_r-不等式). 设 $r > 0$, 则

$$E(|\xi + \eta|^r|\mathscr{C}) \leqslant C_r E(|\xi|^r|\mathscr{C}) + C_r E(|\eta|^r|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.},$$

其中

$$C_r = \begin{cases} 1, & r \leqslant 1, \\ 2^{r-1}, & r \geqslant 1. \end{cases}$$

VIII) (Hölder 不等式). 若 $r > 1, s > 1, \dfrac{1}{r} + \dfrac{1}{s} = 1$, 则

$$E(|\xi\eta||\mathscr{C}) \leqslant [E(|\xi|^r|\mathscr{C})]^{\frac{1}{r}} \cdot [E(|\eta|^s|\mathscr{C})]^{\frac{1}{s}}, P_{\mathscr{C}}\text{-a.e.},$$

IX)(Minkowski 不等式). 若 $r \geqslant 1$, 则

$$\begin{aligned} [E(|\xi + \eta|^r|\mathscr{C})]^{\frac{1}{r}} \leqslant & [E(|\xi|^r|\mathscr{C})]^{\frac{1}{r}} \\ & + [E(|\eta|^r|\mathscr{C})]^{\frac{1}{r}}, P_{\mathscr{C}}\text{-a.e.} \end{aligned}$$

X) 若 $r > 0$, 则 $[E(|\xi|^r|\mathscr{C})]^{\frac{1}{r}}, P_{\mathscr{C}}$-a.e. 不降.

以上各不等式中等号 $P_{\mathscr{C}}$-a.e. 成立的条件与第 3 章 3.6.1 所述相同.

证　i) C_r 不等式的证明如下：对一切 $B \in \mathscr{C}$,

$$\begin{aligned} \int_B E(|\xi + \eta|^r|\mathscr{C})dP &= \int_B |\xi + \eta|^r dP \\ &\leqslant \int_B (C_r|\xi|^r + C_r|\eta|^r)dP \\ &= \int_B [C_r E(|\xi|^r|\mathscr{C}) + C_r E(|\eta|^r|\mathscr{C})]dP. \end{aligned}$$

这里用到了条件期望的定义、关于数列的 C_r-不等式、积分性质以及 II.

ii) Hölder 不等式的证明: 先证 ξ,η 为有界函数的情形. 若记集 $B_1 \hat{=} \{E(|\xi|^r|\mathscr{C}) = 0\}, B_2 \hat{=} \{E(|\eta|^s|\mathscr{C}) = 0\}$, 于是

$$\int_{B_1} |\xi|^r dP = \int_{B_1} E(|\xi|^r|\mathscr{C})dP = 0,$$

即

$$\chi_{B_1}|\xi|^r = 0, \quad P\text{-a.e.}$$

同理可证

$$\chi_{B_2}|\eta|^s = 0, \quad P\text{-a.e.}$$

于是

$$|\xi\eta|\chi_{B_1\cup B_2} = 0, \quad P\text{-a.e.}$$

再应用 5.2.3 条件期望的平滑性质及 $B_1 \cup B_2$ 是 $\mathscr{C}$ 可测的, 知

$$E(\chi_{B_1\cup B_2}|\xi\eta||\mathscr{C}) = \chi_{B_1\cup B_2}E(|\xi\eta||\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}$$

亦即在 $B_1 \cup B_2$ 上, $E(|\xi\eta||\mathscr{C}) = 0, P_{\mathscr{C}}$-a.e. 因而 Hölder 不等式在 $B_1 \cup B_2$ 上自然成立.

在 $B_1^c \cap B_2^c$ 上, $E(|\xi|^r|\mathscr{C}) \neq 0, E(|\eta|^s|\mathscr{C}) \neq 0$, 在不等式

$$a^\alpha b^\beta \leqslant \alpha a + \beta b, (a \geqslant 0, b \geqslant 0, 0 < \alpha < 1, \alpha + \beta = 1)$$

中, 令

$$a = \frac{|\xi|^r}{E(|\xi|^r|\mathscr{C})}, b = \frac{|\eta|^s}{E(|\eta|^s|\mathscr{C})}, \alpha = \frac{1}{r}, \beta = \frac{1}{s},$$

即得

$$\frac{|\xi\eta|}{E^{\frac{1}{r}}(|\xi|^r|\mathscr{C})E^{\frac{1}{s}}(|\eta|^r|\mathscr{C})} \leqslant \frac{1}{r}\frac{|\xi|^r}{E(|\xi|^r|\mathscr{C})} + \frac{1}{s}\frac{|\eta|^s}{E(|\eta|^s|\mathscr{C})}.$$

两边乘以 $\chi_{B_1^c\cap B_2^c}$, 取 $\mathscr{C}$ 之下的条件期望, 注意 $E^{\frac{1}{r}}(|\xi|^r|\mathscr{C}), E^{\frac{1}{s}}(|\eta|^s|\mathscr{C})$ 及 $\chi_{B_1^c\cap B_2^c}$ 都是 $\mathscr{C}$-可测的, 应用 5.2.3 条件期望的平滑性质以及 III 即得

$$\chi_{B_1^c\cap B_2^c}\frac{E(|\xi\eta||\mathscr{C})}{E^{\frac{1}{r}}(|\xi|^r|\mathscr{C})E^{\frac{1}{s}}(|\eta|^s|\mathscr{C})} \leqslant \chi_{B_1^c\cap B_2^c}, P_{\mathscr{C}}\text{-a.e.}$$

即在 $B_1^c \cap B_2^c$ 上仍有 Hölder 不等式成立. 因此当 ξ,η 有界时 Hölder 不等式获证. 对于一般情形, 可按照证明定理 1 中 III 的方法 (截割法), 通过截割及取极限的方法证明.

iii) Minkowski 不等式当 $r = 1$ 时显然成立, 当 $r > 1$ 时, 利用 Hölder 不等式,

注意到 $(r-1)s=r$ 及 $1-\frac{1}{r}=\frac{1}{s}$, 则有

$$\begin{aligned}E(|\xi+\eta|^r|\mathscr{C}) &\leqslant E(|\xi||\xi+\eta|^{r-1}|\mathscr{C})+E(|\eta||\xi+\eta|^{r-1}|\mathscr{C})\\&\leqslant E^{\frac{1}{r}}(|\xi|^r|\mathscr{C})E^{(1-\frac{1}{r})}(|\xi+\eta|^r|\mathscr{C})\\&\quad+E^{\frac{1}{r}}(\eta|^r|\mathscr{C})E^{(1-\frac{1}{r})}(|\xi+\eta|^r|\mathscr{C})\end{aligned}$$

$P_{\mathscr{C}}$-a.e. 成立. 当 $E(|\xi+\eta|^r|\mathscr{C})(\omega)=0$ 时, Minkowski 不等式自动成立; 当 $E(|\xi+\eta|^r|\mathscr{C})(\omega)\neq 0$ 时, 在上述不等式两端同时除以 $E^{(1-\frac{1}{r})}(|\xi+\eta|^r|\mathscr{C})(\omega)$, 即得 Minkowski 不等式.

iv) 设 $r'<r$, 在关于条件期望的 Hölder 不等式中取 r 为 $\frac{r}{r'}$, 取 ξ 为 $|\xi|^{r'},\eta=1$, 即得

$$E(|\xi|^{r'}|\mathscr{C})\leqslant E(|\xi|^r|\mathscr{C})^{\frac{r'}{r}},P_{\mathscr{C}}\text{-a.e.}$$

此即

$$E^{\frac{1}{r'}}(|\xi|^{r'}|\mathscr{C})\leqslant E^{\frac{1}{r}}(|\xi|^r|\mathscr{C}),P_{\mathscr{C}}\text{-a.e.}$$

即 X 获证. □

定理 4　若 $\xi_n\xrightarrow{r}\xi,r\geqslant 1$, 则

$$E(\xi_n|\mathscr{C})\xrightarrow{r}E(\xi|\mathscr{C}).$$

证　由 $\xi_n\xrightarrow{r}\xi,r\geqslant 1$, 知 $E|\xi_n|,E|\xi|$ 有限, 因而

$$\begin{aligned}|E(\xi_n|\mathscr{C})-E(\xi|\mathscr{C})|&=|E(\xi_n-\xi|\mathscr{C})|\\&\leqslant E(|\xi_n-\xi||\mathscr{C})\leqslant E^{\frac{1}{r}}(|\xi_n-\xi|^r|\mathscr{C}),P_{\mathscr{C}}-\text{a.e.}\end{aligned}$$

故

$$\begin{aligned}E|E(\xi_n|\mathscr{C})-E(\xi|\mathscr{C})|^r&\leqslant E(E(|\xi_n-\xi|^r|\mathscr{C}))\\&=E|\xi_n-\xi|^r\to 0.\end{aligned}$$

因而定理获证. □

5.2.3　给定 σ- 代数下条件期望的平滑性质

简单地说, 给定 σ-代数 $\mathscr{C}$ 下的条件期望是一个 $\mathscr{C}$-平滑算子, 在一定意义下它具有数学期望的取平均值的功能. 和此有关的一类性质总称为平滑性质.

我们已经知道, $E(\xi|\mathscr{C})$ 是一个 $\mathscr{C}$-可测函数, 因而它在 $\mathscr{C}$ 的每一个原子上取同一个值. 而且我们还将看到在 $\mathscr{C}$ 的原子上它取 ξ 关于概率 P 的平均值, 即

定理 5 在 $\mathscr{C}$ 的每个原子 B 上, $E(\xi|\mathscr{C})$ 是常数. 且若 $P(B)>0$, 则对一切 $\omega\in B$

$$E(\xi|\mathscr{C})(\omega)=\frac{1}{P(B)}\int_B \xi dP, \tag{4}$$

记此常数为 $E(\xi|B)$.

证 由于 $P(B)>0$, 故有一 $\omega_0\in B$, 则由 $E(\xi|\mathscr{C})$ 是 $\mathscr{C}$-可测函数知

$$B_0\hat{=}\{\omega: E(\xi|\mathscr{C})(\omega)=E(\xi|\mathscr{C})(\omega_0)\}\cap B\in\mathscr{C},$$

由于 $B_0\subset B$, 且 $B_0\neq\varnothing$. 因而由原子的定义 (除本身及空集 $\varnothing$ 而外不包含任何可测子集), $B_0=B$ 即 $E(\xi|\mathscr{C})$ 在 B 上的值是常数, 记为 $E(\xi|B)$, 则

$$E(\xi|B)P(B)=\int_B E(\xi|\mathscr{C})dP=\int_B \xi dP.$$

因 $P(B)>0$, 故 (4) 式成立.

这个定理清楚地说明 $E(\xi|\mathscr{C})$ 在 $\mathscr{C}$ 的原子但非 $\mathscr{A}$ 的原子上是"ξ 的平均值", 在这个意义上, $E(\xi|\mathscr{C})$ 是 ξ 的一个 $\mathscr{C}$-平滑函数.

由此定理还可得到 §5.1 曾经提到的结果.

推论 1 设 $\mathscr{C}$ 是由 Ω 的可数分割 $\{B_n\}\subset\mathscr{A}$ 生成的最小 σ-代数, 则

$$E(\xi|\mathscr{C})=\sum_n E(\xi|B_n)\chi_{B_n},$$

其中当 $P(B_n)\neq 0$ 时,$E(\xi|B_n)$ 由 (4) 式的右端定义, 而当 $P(B_n)=0$ 时,$E(\xi|B_n)$ 取任一常数.

其次, 对于 $\mathscr{A}$ 的最小子 σ-代数 $\mathscr{C}_0=\{\varnothing,\Omega\}$, 则有

$$E(\xi|\mathscr{C}_0)=E\xi.$$

这一结论对每一与 $\sigma(\xi)$(即 $\xi^{-1}(\mathscr{B}^{(1)})$ 或 $\xi^{-1}(\mathscr{B}_z^{(1)})$) 独立的 σ-代数 $\mathscr{C}$ 也是 $P_{\mathscr{C}}$-a.e. 成立的. 即

定理 6 若 $\mathscr{C}$ 与 $\sigma(\xi)$ 独立, 则

$$E(\xi|\mathscr{C})=E\xi,\quad P_{\mathscr{C}}\text{-a.e.}$$

证 首先, 当 $\xi\geqslant 0$ 时定理成立, 因为对每一 $B\in\mathscr{C}$ 来说, χ_B 与 ξ 独立, 利用乘法定理 (第 3 章, §3.4.1, 性质 3),

$$\int_B E(\xi|\mathscr{C})dP=\int_B \xi dP=E\xi\chi_B=E\xi\cdot P(B)=\int_B E\xi dP.$$

当 ξ 为实可测函数时, 由假设知 $\sigma(\xi^+),\sigma(\xi^-)$ 分别与 $\mathscr{C}$ 独立, 所以,$E(\xi|\mathscr{C})=E(\xi^+|\mathscr{C})-E(\xi^-|\mathscr{C})=E\xi^+-E\xi^-=E\xi$, a.e.$P_{\mathscr{C}}$. 当 ξ 为复可测函数时, 可通过实部和虚部类似地证明. □

再由 §2.1 性质 2 知当 ξ 是 $\mathscr{C}$-可测函数时,$E(\xi|\mathscr{C})=\xi,P_{\mathscr{C}}$-a.e.. 根据前面的解释, 可以认为 $E(\cdot|\mathscr{C})$ 对 $\mathscr{C}$-可测函数不再起平滑作用. 更一般地, 有下列的

定理 7　若 ξ 是 $\mathscr{C}$-可测的,η 是使 $E\xi\eta,E\eta$ 存在的随机变量且 ξ,η 之一是实的, 或 $\xi\eta,\eta$ 是可积的复随机变量, 则

$$E(\xi\eta|\mathscr{C})=\xi E(\eta|\mathscr{C}),P_{\mathscr{C}}\text{-a.e..}\tag{5}$$

证　i) 先考虑 $\eta\geqslant 0$ 的情形. 首先, 当 $\xi=\chi_{B'},B'\in\mathscr{C}$ 时, 对任何 $B\in\mathscr{C}$,

$$\begin{aligned}\int_B E(\chi_{B'}\cdot\eta|\mathscr{C})dP&=\int_B\chi_{B'}\eta dP\\&=\int_{B\cap B'}E(\eta|\mathscr{C})dP=\int_B\chi_{B'}E(\eta|\mathscr{C})dP.\end{aligned}$$

因而 (5) 对 $\eta\geqslant 0,\xi=\chi_{B'},B'\in\mathscr{C}$ 成立, 因而也对 ξ 是 $\mathscr{C}$-可测的非负简单函数成立. 由条件期望的单调收敛性知 (5) 对 ξ 是 $\mathscr{C}$-可测的非负函数成立. 当 ξ 为 $\mathscr{C}$-可测的实函数时, 由 $E\xi\eta$ 存在知 $E\xi^+\eta,E\xi^-\eta$ 及 $E\xi^+\eta-E\xi^-\eta$ 存在, 因而

$$\begin{aligned}E(\xi\eta|\mathscr{C})&=E(\xi^+\eta|\mathscr{C})-E(\xi^-\eta|\mathscr{C})\\&=\xi^+E(\eta|\mathscr{C})-\xi^-E(\eta|\mathscr{C})\\&=\xi E(\eta|\mathscr{C}),P_{\mathscr{C}}\text{-a.e.}\end{aligned}$$

对于 ξ 是复可测的情形, 设 $\xi=\xi_1+i\xi_2,\xi_1,\xi_2$ 为实的 $\mathscr{C}$-可测函数. 由 $E\xi\eta$ 存在知 $E\xi_1\eta+iE\xi_2\eta$ 存在, 由定理 1,II, 及上述结果知

$$\begin{aligned}E(\xi\eta|\mathscr{C})&=E(\xi_1\eta|\mathscr{C})+iE(\xi_2\eta|\mathscr{C})\\&=\xi_1E(\eta|\mathscr{C})+i\xi_2E(\eta|\mathscr{C})\\&=\xi E(\eta|\mathscr{C}),P_{\mathscr{C}}\text{-a.e..}\end{aligned}$$

ii) 再考虑 η 为实随机变量的情形. 若 ξ 为实的 $\mathscr{C}$-可测函数, 由 $E\xi\eta$ 存在, 知 $E(\xi^+\eta^++\xi^-\eta^-),E(\xi^+\eta^-+\xi^-\eta^+)$ 有一有限, 因而 $E\xi\eta^+,E\xi\eta^-$ 及 $E\xi\eta^+-E\xi\eta^-$ 有意义, 于是根据定理 1 的 II 及 i) 中结果, 有

$$\begin{aligned}E(\xi\eta|\mathscr{C})&=E(\xi\eta^+-\xi\eta^-|\mathscr{C})=E(\xi\eta^+|\mathscr{C})-E(\xi\eta^-|\mathscr{C})\\&=\xi E(\eta^+|\mathscr{C})-\xi E(\eta^-|\mathscr{C})=\xi(E(\eta^+|\mathscr{C})-E(\eta^-|\mathscr{C}))\\&=\xi E(\eta|\mathscr{C}),P_{\mathscr{C}}\text{-a.e.}\end{aligned}$$

同 i) 中证明 ξ 是复的 $\mathscr{C}$-可测函数时的方法一样可证 ξ 是复的 $\mathscr{C}$-可测函数,η 为实随机变量时 (5) 成立.

iii) η 是复随机变量, ξ 是实的 $\mathscr{C}$-可测函数时, 设 $\eta=\eta_1+i\eta_2,\eta_1,\eta_2$ 为实随机变量. 由于 $E\xi\eta$ 存在意味着 $E\xi\eta_1,E\xi\eta_2$ 及 $E\xi\eta_1+iE\xi\eta_2$ 存在, 利用 ii) 中结果及定理 1 的 II, 易知在此条件下 (5) 式成立.

iv) 当 ξ,η 均为有界复随机变量时, 设 $\xi=\xi_1+i\xi_2,\eta=\eta_1+i\eta_2,\xi_1,\xi_2,\eta_1,\eta_2$ 皆为有界实随机变量, 此时 $E\xi\eta=E\xi_1\eta_1-E\xi_2\eta_2+i(E\xi_1\eta_2+E\xi_2\eta_1)$, 因而由 ii) 中结果及条件期望的线性性质知 (5) 成立.

再利用条件期望的控制收敛性及截割法可证当 $\xi\eta$ 及 η 可积时 (5) 式成立.

定理所述的各种情况皆已证明. □

定理 8 若 $\mathscr{C}\subset\mathscr{C}'$, 则

$$E(E(\xi|\mathscr{C})|\mathscr{C}')=E(\xi|\mathscr{C})=E(E(\xi|\mathscr{C}')|\mathscr{C}),P_{\mathscr{C}}\text{-a.e.} \tag{6}$$

证 由于 $E(\xi|\mathscr{C})$ 是 $\mathscr{C}$-可测函数, 因而由 $\mathscr{C}\subset\mathscr{C}'$ 知它是 $\mathscr{C}'$ 可测的, 由于 $E(E(\xi|\mathscr{C}))=E\xi$ 存在, 由定理 7 知

$$E(E(\xi|\mathscr{C})|\mathscr{C}')=E(\xi|\mathscr{C})E(1|\mathscr{C}')=E(\xi|\mathscr{C}),P_{\mathscr{C}}\text{-a.e.}$$

此即 (6) 的第一个等式. 至于第二个等式则因对任一 $B\in\mathscr{C}$(因而 $B\in\mathscr{C}'$) 来说, 由定义 1 知

$$\begin{aligned}\int_B E(E(\xi|\mathscr{C}')|\mathscr{C})dP&=\int_B E(\xi|\mathscr{C}')dP\\&=\int_B\xi dP=\int_B E(\xi|\mathscr{C})dP.\end{aligned}$$

于是即得 (6) 的第二等式. □

注 从定理 8 我们看到扩展条件期望定义的必要性. 在定理 8 中即令 ξ 是随机变量,$E(\xi|\mathscr{C})$ 也未必是随机变量, 因而 $E(E(\xi|\mathscr{C})|\mathscr{C}')$ 就不是随机变量的条件期望, 而是一个可测函数的条件期望. 在证明定理 8 的过程中应用了定理 7, 且定理 7 中的 ξ 不一定是随机变量, 因而 $E(\xi\eta|\mathscr{C})$ 也是一个一般可测函数的条件期望. 请读者结合定义 1 后的注 1° 和 2° 进一步体会.

最后, 我们应用上面的定理来证明

定理 9 若 $\xi\in L_2(\Omega,\mathscr{A},P),\mathscr{C}$ 是 $\mathscr{A}$ 的子 σ-代数, 则 $E(\xi|\mathscr{C})\in L_2(\Omega,\mathscr{C},P_{\mathscr{C}})$, 且 $E(\xi|\mathscr{C})$ 是 ξ 在 $L_2(\Omega,\mathscr{C},P_{\mathscr{C}})$ 中的**最佳均方逼近**, 即对任一 $f\in L_2(\Omega,\mathscr{C},P_{\mathscr{C}})$ 来说,

$$E|\xi-E(\xi|\mathscr{C})|^2\leqslant E|\xi-f|^2 \tag{7}$$

并且进一步有

$$E(|\xi - E(\xi|\mathscr{C})|^2|\mathscr{C}) \leqslant E(|\xi - f|^2|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.} \tag{8}$$

证　只需证后一结论, 因为对 (8) 取数学期望即得 (7).

由定理 1 的 III 及定理 3 的 X 知

$$|E(\xi|\mathscr{C})|^2 \leqslant |E(|\xi||\mathscr{C})|^2 \leqslant E(|\xi|^2|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}$$

而 $E(E(|\xi|^2|\mathscr{C})) = E|\xi|^2 < \infty$, 故 $E(\xi|\mathscr{C})$ 平方可积, 且 $E(\xi|\mathscr{C})$ 为 $\mathscr{C}$-可测函数, 因而 $E(\xi|\mathscr{C}) \in L_2(\Omega, \mathscr{C}, P_{\mathscr{C}})$.

其次, 对任一 $f \in L_2(\Omega, \mathscr{C}, P_{\mathscr{C}})$, 设 ξ, f 是复的, 而 $\xi = \xi_1 + i\xi_2, f = f_1 + if_2, \xi_1, \xi_2, f_1, f_2$ 是实的, 则

$$\begin{aligned} E(|\xi_i - f_i|^2|\mathscr{C}) &= E[|\xi_i - E(\xi_i|\mathscr{C}) + E(\xi_i|\mathscr{C}) - f_i|^2|\mathscr{C}] \\ &= E[|\xi_i - E(\xi_i|\mathscr{C})|^2|\mathscr{C}] + E[|E(\xi_i|\mathscr{C}) - f_i|^2|\mathscr{C}] \\ &\quad + 2E[(\xi_i - E(\xi_i|\mathscr{C}))(E(\xi_i|\mathscr{C}) - f_i)|\mathscr{C}], \\ &\quad P_{\mathscr{C}}\text{-a.e.} i = 1, 2. \end{aligned}$$

由于 $E(\xi_i|\mathscr{C}) - f_i$ 是 $\mathscr{C}$ 可测函数, 故由定理 7 知上式右端的第三项为

$$[E(\xi_i|\mathscr{C}) - f_i]E[\xi_i - E(\xi_i|\mathscr{C})|\mathscr{C}] = 0, P_{\mathscr{C}}\text{-a.e.} i = 1, 2.$$

因而

$$\begin{aligned} E(|\xi_i - f_i|^2|\mathscr{C}) &= E(|\xi_i - E(\xi_i|\mathscr{C})|^2|\mathscr{C}) \\ &\quad + E(|E(\xi_i|\mathscr{C}) - f_i|^2|\mathscr{C}) \geqslant E(|\xi_i - E(\xi_i|\mathscr{C})|^2|\mathscr{C}), \\ &\quad P_{\mathscr{C}}\text{-a.e.} i = 1, 2. \end{aligned}$$

由于 $|\xi - f|^2 = |\xi_1 - f_1|^2 + |\xi_2 - f_2|^2, |\xi - E(\xi|\mathscr{C})|^2 = |\xi_1 - E(\xi_1|\mathscr{C})|^2 + |\xi_2 - E(\xi_2|\mathscr{C})|^2$, 所以

$$\begin{aligned} E(|\xi - f|^2|\mathscr{C}) &= E(|\xi_1 - f_1|^2|\mathscr{C}) + E(|\xi_2 - f_2|^2|\mathscr{C}) \\ &\geqslant E(|\xi_1 - E(\xi_1|\mathscr{C})|^2|\mathscr{C}) + E(|\xi_2 - E(\xi_2|\mathscr{C})|^2|\mathscr{C}) \\ &= E(|\xi - E(\xi|\mathscr{C})|^2|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.} \end{aligned}$$

定理获证. □

实际上, 常常遇见这样的问题: 设 $\xi, \xi_1, \cdots, \xi_n$ 是二阶矩存在的随机变量. 希望求出 $\xi_1, \cdots, \xi_n$ 的一个函数 $f(\xi_1, \cdots, \xi_n)$ 来 "最好地逼近" ξ(通常称为回归问题). 若 "最好地逼近" 的意义理解为 "最佳均方逼近", 则由定理 9 及 §5.3 定理 1 知

$$f(\xi_1, \cdots, \xi_n) = E(\xi|\sigma(\xi_1, \cdots, \xi_n))$$

即为所求. 但是也应同时指出: 除了某些特殊情形 (如按正态律分布的情形) 外, 还没有很好的实际可行的方法来求函数 $f(x_1,\cdots,x_n)$. 如果 $(\xi,\xi_1,\cdots,\xi_n)$ 按正态律分布, $f(x_1,\cdots,x_n)$ 的形式将在 §3 例 7 中给出.

为了应用, 我们退一步将问题改成下面的形式: 设 $\xi,\xi_1,\cdots,\xi_n$ 是二阶矩存在的随机变量, 希望求 $\xi_1,\cdots,\xi_n$ 的一个线性组合 $\sum\limits_{k=1}^{n}a_k\xi_k$, 使得对任何 $b_1,\cdots,b_n$,

$$E|\xi-\sum_{k=1}^{n}a_k\xi_k|^2\leqslant E|\xi-\sum_{k=1}^{n}b_k\xi_k|^2. \tag{9}$$

这就是所谓的线性回归问题. 当 $\xi,\xi_1,\cdots,\xi_n$ 为实随机变量时, 线性回归问题不难由古典分析的方法解决 (参看 [9] 第 33 章). 为了便于与定理 9 的结论对比, 我们有时称满足 (9) 的线性函数 $\sum\limits_{k=1}^{n}a_k\xi_k$ 为 ξ 在 $(\xi_1,\cdots,\xi_n)$ 下的广义条件数学期望, 而记作 $\hat{E}(\xi|\xi_1,\cdots,\xi_n)$. 更一般地可以令 $\mathfrak{M}$ 为 $L_2(\Omega,\mathscr{A},P)$ 的一个线性子空间, $\xi\in L_2(\Omega,\mathscr{A},P)$**在 $\mathfrak{M}$ 下的广义条件数学期望**是指满足条件

$$E|\xi-\hat{E}(\xi|\mathfrak{M})|^2\leqslant E(|\xi-f|^2),$$

对任何 $f\in\mathfrak{M}$ 的 $\mathfrak{M}$ 中的函数 $\hat{E}(\xi|\mathfrak{M})$. 广义条件期望与条件期望有着很多相似的性质, 不过我们不再在此停留了.

习题及补充

1. 在条件期望的控制收敛性 VI 中, 其他条件均成立, 将 $\xi_n\to\xi, P$-a.e. 改成 $\xi_n\xrightarrow{P}\xi$, 结果如何?

2. 若 g 为 R 上连续凸函数, 且 $E\xi, Eg(\xi)$ 有限,$E(\xi|\mathscr{C})$ 有限, 则

$$g(E(\xi|\mathscr{C}))\leqslant E(g(\xi)|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}$$

3. i) 若 $\mathscr{C}\subset\mathscr{C}'\subset\mathscr{A}$, 且 ξ' 是 $\mathscr{C}'$ 可测的,$E\xi\xi'$,$E\xi$ 存在 (若 ξ,ξ' 皆为复函数时, 要求 $\xi\xi',\xi$ 可积), 则

$$E(\xi\xi'|\mathscr{C})=E(\xi' E(\xi|\mathscr{C}')|\mathscr{C}), P_{\mathscr{C}}\text{-a.e.}$$

ii) 若 i) 成立, 则 5.2.3 中的定理 7,8 成立.

4. 设 σ-代数 $\mathscr{A}_n\subset\mathscr{A}$, 且 $\mathscr{A}_n\uparrow,\xi_n$ 是 $\mathscr{A}_n$ 可测的, $n=1,2,\cdots$, 若对一切 n 及 $m\leqslant n$,

$$E(\xi_n|\mathscr{A}_m)=\xi_m(\geqslant\xi_m,\leqslant\xi_m), \text{a.e.}$$

则称 $\{\xi_n\}$ 为一鞅 (相应地下鞅、上鞅).

若 $\eta_1,\cdots,\eta_n,\cdots$ 为一独立实随机变量序列, 若对一切 $k,n,E\eta_k=0(E\eta_k\geqslant 0),\mathscr{A}_n=\sigma(\eta_1,\cdots,\eta_n),\xi_n=\sum_{k=1}^{n}\eta_k$, 试证 $\{\xi_n\}$ 为一鞅 (相应地下鞅).

§5.3　给定函数下的条件数学期望

在初等概率论中, 也曾遇到过给定随机变量下条件数学期望的问题. 例如 (ξ,η) 为二维连续型随机变量, 要考虑 $E(\xi|\eta=y),y\in R^{(1)}$. 由于 $P(\eta=y)=0$, 因此, 它不能按照给定事件下的条件期望来定义. 在一些教科书中采取先定义条件分布函数

$$F(x|y)\hat{=}P(\xi<x|\eta=y)\hat{=}\lim_{\Delta y\to 0}P(\xi<x|y\leqslant\eta<y+\Delta y),$$

然后再定义

$$E(\xi|\eta=y)\hat{=}\int xdF(x|y).$$

但这样做法要想达到数学上的严格性会遭遇到古典分析的困难.

在这一节, 我们用测度论的工具讨论在给定映射下随机变量的条件期望及其性质. 作为特例我们将讨论在给定随机函数 $\eta_T=\{\eta_t,t\in T\}$ 下随机变量 ξ 的条件期望. 最后, 对于连续型随机变量的情形, 得出与通常教科书中一致的结论.

设 ξ,η 为 $(\Omega,\mathscr{A},P)$ 上的随机变量, $E\xi$ 存在, ξ 在给定 σ-代数 $\sigma(\eta)$ 下的条件期望为 $E(\xi|\sigma(\eta))$, 由于 $\{\eta=y\}$ 是 $\sigma(\eta)$ 的原子, 因而在 $\{\eta=y\}$ 上 $E(\xi|\sigma(\eta))$ 为常数. 于是 $E(\xi|\sigma(\eta))$ 可以看作是 η 的函数, 记作 $E(\xi|\eta)$, 称为 ξ 在给定随机变量 η 下的条件数学期望.

为了进一步推广这一概念, 我们先来回顾第 2 章讨论过的可测映射的概念.

设 $(\Omega,\mathscr{A})$ 是一可测空间,f 是由 Ω 到 Ω' 的映射, 则可以证明存在着一个 Ω' 的子集作成的 σ-代数 $\mathscr{A}'$, 使得 f 是 $(\Omega,\mathscr{A})$ 到 $(\Omega',\mathscr{A}')$ 的可测映射. 事实上, 令

$$\mathscr{A}'=\{A':A'\subset\Omega',f^{-1}(A')\in\mathscr{A}\}.\tag{1}$$

则仿照第 2 章 §2.2 证明性质 2 的方法, 可以证明 $\mathscr{A}'$ 是 Ω' 中的 σ-代数. 于是显然 $\mathscr{A}'$ 满足上述要求.

定理 1　设 ξ 是 $(\Omega,\mathscr{A},P)$ 上的一个随机变量, $E\xi$ 存在, f 是 $(\Omega,\mathscr{A})$ 到 $(\Omega',\mathscr{A}')$ 中的可测映射, 则

$$E(\xi|\sigma(f))=g\circ f,P_{\sigma(f)}\text{-a.e.}\tag{2}$$

其中 g 是满足

$$\int_{A'}gdP'_f=\int_{f^{-1}(A')}\xi dP,A'\in\mathscr{A}'\tag{3}$$

而由 $E(\xi|\sigma(f)), P'_f$-a.e. 唯一决定的 $(\Omega', \mathscr{A}')$ 上的一个可测函数. 其中 P'_f 是 P 在 $(\Omega', \mathscr{A}')$ 上由 f 导出的概率测度. (参看第 3 章 §3.4 定义 6.)

证 设 $\xi = \xi_1 + i\xi_2, \xi_1, \xi_2$ 是实随机变量, 则

$$E(\xi|\sigma(f)) = E(\xi_1|\sigma(f)) + iE(\xi_2|\sigma(f)).$$

由于 $E(\xi_i|\sigma(f))$ 是 $\sigma(f)$ 可测的实函数, 由第 2 章 §2.2 定理 13 知存在 $(\Omega', \mathscr{A}')$ 上实可测函数 g_j, 使得

$$E(\xi_j|\sigma(f)) = g_j \circ f, j = 1, 2.$$

令

$$g = g_1 + ig_2,$$

则

$$E(\xi|\sigma(f)) = g_1 \circ f + ig_2 \circ f = g \circ f,$$

即 (2) 成立.

再由第 3 章 §3.4 积分变换定理、(2) 式及条件期望的定义得

$$\begin{aligned}\int_{A'} g dP'_f &= \int_{f^{-1}(A')} g \circ f dP \\ &= \int_{f^{-1}(A')} E(\xi|\sigma(f)) dP = \int_{f^{-1}(A')} \xi dP, A' \in \mathscr{A}',\end{aligned}$$

即 (3) 成立.

最后, 令

$$\varphi(A') = \int_{f^{-1}(A')} E(\xi|\sigma(f)) dP = \int_{A'} g(\omega') dP'_f, A' \in \mathscr{A}',$$

则 φ 是 $\mathscr{A}'$ 上 σ-可加,P'_f-连续集函数, 因而 g 由 $E(\xi|\sigma(f)), P'_f$-a.e. 唯一决定. □

由于 $E(\xi|\sigma(f))$ 在 $\{f = \omega'\} (\omega' \in \Omega')$ 上为常数,(为什么?) 因而常以 $E(\xi|f)$ 表示 $E(\xi|\sigma(f))$, 又由于当 $f(\omega) = \omega'$ 时

$$E(\xi|f)(\omega) = g(f(\omega)) = g(\omega'),$$

除去 $\sigma(f)$ 中的零概率集 (因而除去 $\mathscr{A}'$ 中的零概率集) 而外, 对一切 $\omega \in \Omega$ 成立, 因而我们也可以认为 ξ 在 f 之下的条件期望是定理 1 中的 g, 并记作

$$E(\xi|f = \omega') \widehat{=} g(\omega'), \omega' \in \Omega'. \tag{4}$$

例 1 当 f 为 $(\Omega, \mathscr{A}, P)$ 上的实 (或复) 随机变量 η 时,

$$E(\xi|\eta = y) = g(y), y \in R^{(1)} (\text{相应地}, y \in Z^{(1)}),$$

其中 g 是 $(R^{(1)},\mathscr{B}^{(1)})$(相应地,$(Z^{(1)},\mathscr{B}_z^{(1)})$) 上的可测函数.

例 2　当 f 是 $(\Omega,\mathscr{A},P)$ 上的 m 维实 (或复) 随机变量 $(\eta_1,\cdots,\eta_m)$ 时,

$$E(\xi|\eta_1=y_1,\cdots,\eta_m=y_m)=g(y_1,\cdots,y_m), s$$

$$(y_1,\cdots,y_m)\in R^{(m)}(\text{或}Z^{(m)}),$$

其中 g 是 $(R^{(m)},\mathscr{B}^{(m)})$(相应地,$(Z^{(m)},\mathscr{B}_z^{(m)})$) 上的可测函数.

以上两例中的 g 均按定理 1 决定.

当给定的 f 是随机变量族的情形, 在研究马氏过程时很有用, 在此作些讨论.

第 4 章已经给出了无穷乘积空间和无穷乘积 σ-代数的定义. 设 T 是任一指标集,

$$R^T=\prod_{t\in T}R_t,\quad Z^T=\prod_{t\in T}Z_t,$$

$$\mathscr{B}^T=\prod_{t\in T}\mathscr{B}_t=\bigcup_{T_c}\mathscr{B}^{T_c}\times\prod_{t\in T-T_c}R_t,$$

$$\mathscr{B}_z^T=\prod_{t\in T}\mathscr{B}_{z_t}=\bigcup_{T_c}\mathscr{B}_z^{T_c}\times\prod_{t\in T-T_c}Z_t,$$

其中 T_c 跑遍 T 的一切可数子集且对一切 $t\in T,R_t=R^{(1)},Z_t=Z^{(1)},\mathscr{B}_t=\mathscr{B}^{(1)},\mathscr{B}_{z_t}=\mathscr{B}_z^{(1)}$. 则 $(R^T,\mathscr{B}^T),(Z^T,\mathscr{B}_z^T)$ 为可测空间.

定义 1　设 $(\Omega,\mathscr{A},P)$ 为一概率空间, 称 $(\Omega,\mathscr{A})$ 到 $(R^T,\mathscr{B}^T)$(或 $(Z^T,\mathscr{B}_z^T)$) 的可测映射 $\eta_{_T}=\{\eta_t,t\in T\}$ 为**随机函数**. η_t 称为**随机函数** $\eta_{_T}$ **的分量**

一个随机函数 $\eta_{_T}$, 给定 $\omega\in\Omega$, 是 $t\in T$ 上的函数; 给定 $t\in T,\eta_t$ 是一个随机变量. 由乘积 σ-代数的定义及第 2 章 2.2.1 性质 3 易证 $\eta_{_T}$ 为 $(\Omega,\mathscr{A},P)$ 上随机函数的充分必要条件是它的每一分量 η_t 是 $(\Omega,\mathscr{A},P)$ 上的随机变量. 因此 $\eta_{_T}$ 是定义域为 T 而取值为随机变量的函数 (此即随机函数一词的由来).

由第 2 章 2.2.4 定理 13 推知, 若 $\eta_{_T}$ 为实 (复) 随机函数, 则 φ 是 $\sigma(\eta_{_T})$ 可测函数的充分必要条件是存在 $(R^T,\mathscr{B}^T)$(相应地,$(Z^T,\mathscr{B}_z^T)$) 上的可测函数 g, 使 $\varphi(\omega)=(g\circ\eta_{_T})(\omega)=g(\eta_T(\omega))$, 对一切 $\omega\in\Omega$ 成立. 因而有

例 3　设 ξ 是 $(\Omega,\mathscr{A},P)$ 上的随机变量, $E\xi$ 存在, $\eta_{_T}$ 是 $(\Omega,\mathscr{A},P)$ 上的随机函数, 则

$$E(\xi|\eta_{_T})=g\circ\eta_{_T}$$

或

$$E(\xi|\eta_{_T}=y_{_T})=g(y_{_T})$$

称为 ξ 在给定随机函数 η_T 下的条件数学期望. 其中 g 是 $(R^T,\mathscr{B}^T)$(相应地, 当 η_T 是复随机函数时为 $(Z^T,\mathscr{B}_z^T)$) 上的可测函数, 由 $E(\xi|\sigma(\eta_T))$ 通过

$$\int_{B_T} g(y_T)dP_{\eta_T}=\int_{\eta_T^{-1}(B_T)} E(\xi|\sigma(\eta_T))dP, B_T\in\mathscr{B}^T(\text{或}\mathscr{B}_z^T)$$

P_{η_T}-a.e. 唯一决定. 其中 P_{η_T} 为随机函数 η_T 的分布律.

由于条件概率是条件期望的特殊情形, 因此由定理 1 立刻得出关于条件概率的相应结论. 即

推论 1 设 $(\Omega,\mathscr{A},P)$ 是一概率场, $A\in\mathscr{A}, f$ 是 $(\Omega,\mathscr{A})$ 到 $(\Omega',\mathscr{A}')$ 中的一个可测映射, 则 A 在 $\sigma(f)$ 下的条件概率

$$P(A|\sigma(f))=h\circ f, \tag{5}$$

其中 h 是满足

$$\int_{B'} h(\omega')dP'_f=P(A\cap f^{-1}(B')), B'\in\mathscr{A}'$$

而由 $P(A|\sigma(f)), P'_f$-a.e. 唯一决定的 $(\Omega',\mathscr{A}')$ 上的可测函数.

证 只需在定理 1 中令 $\xi=\chi_A$ 即得. □

同定理 1 后面的注解一样, 可以把 $P(A|\sigma(f))$ 记作 $P(A|f)$, 可以把 $h(\omega')$ 记作

$$h(\omega')\widehat{=}P(A|f=\omega'), \omega'\in\Omega'. \tag{6}$$

当 f 是 $(\Omega,\mathscr{A},P)$ 上的随机函数 η_T, 而

$$A=\{\omega:\xi_{T_1}(\omega)\in B_z^{T_1}\},$$

其中 ξ_{T_1} 也是 $(\Omega,\mathscr{A},P)$ 上的随机函数,$B_z^{T_1}\in\mathscr{B}_z^{T_1}$, 则

$$P(A|\eta_T=z_T)=h(z_T), z_T\in Z^T.$$

其中 h 按推论 1 确定.

对于 $\xi_{T_1}=(\xi_1,\cdots,\xi_n), \eta_T=(\eta_1,\cdots,\eta_m)$ 的特殊情形, 我们得到

$$P((\xi_1,\cdots,\xi_n)\in B^{(n)}|\eta_1=z_1,\cdots,\eta_m=z_m),$$

$$(z_1,\cdots,z_m)\in Z^{(m)}$$

的确切含意.

现在我们应用以上的结论来导出通常教科书中关于连续型与离散型随机变量的条件数学期望与条件概率的公式.

例 4　设 $\zeta=(\xi_1,\cdots,\xi_n,\eta_1,\cdots,\eta_m)$ 是 $n+m$ 维连续型实随机变量, 其分布密度为

$$p(x_1,\cdots,x_n,y_1,\cdots,y_m),(x_1,\cdots,x_n,y_1,\cdots,y_m)\in R^{(n+m)},$$

则 $\eta=(\eta_1,\cdots,\eta_m)$ 是 $(\Omega,\mathscr{A})$ 到 $(R^{(m)},\mathscr{B}^{(m)})$ 的可测映射, 今往求条件概率

$$P((\xi_1,\cdots,\xi_n)\in B^{(n)}|\eta_1=y_1,\cdots,\eta_m=y_m),B^{(n)}\in\mathscr{B}^{(n)}$$

由推论 1 及其后面的注解知对任一 $B^{(n)}\in\mathscr{B}^{(n)}$, 记 $\xi=(\xi_1,\cdots,\xi_n)$, 则

$$P(\xi\in B^{(n)}|\eta_1=y_1,\cdots,\eta_m=y_m)=h(y_1,\cdots,y_m), \tag{7}$$

其中 h 满足

$$\underset{B_1^{(m)}}{\int\cdots\int} h(y_1,\cdots,y_m)dP_\eta=P(\xi\in B^{(n)},\eta\in B_1^{(m)}),B_1^{(m)}\in\mathscr{B}^{(m)}. \tag{8}$$

由于 ζ 是连续型随机变量, 因而 η 也是连续型随机变量. 由分布密度的定义及 Fubini 定理知

$$\begin{aligned}&P(\xi\in B^{(n)},\eta\in B_1^{(m)})\\=&\underset{B^{(n)}\times B_1^{(m)}}{\int\cdots\int} p(x_1,\cdots,x_n,y_1,\cdots,y_m)dx_1\cdots dx_ndy_1\cdots dy_m\\=&\underset{B_1^{(m)}}{\int\cdots\int}\left[\underset{B^{(n)}}{\int\cdots\int}p(x_1,\cdots,x_n,y_1,\cdots,y_m)dx_1\cdots dx_n\right]dy_1\cdots dy_m,\end{aligned}$$

而

$$P(\eta\in B_2^{(m)})=\underset{B_2^{(m)}}{\int\cdots\int}p_\eta(y_1,\cdots,y_m)dy_1\cdots dy_m,B_2^{(m)}\in\mathscr{B}^{(m)}.$$

于是由 (8) 及 Radon-Nikodym 定理的推论知, 对任何 $B_1^{(m)}\in\mathscr{B}^{(m)}$ 来说,

$$\begin{aligned}&\underset{B_1^{(m)}}{\int\cdots\int}h(y_1,\cdots,y_m)p_\eta(y_1,\cdots,y_m)dy_1\cdots dy_m\\=&\underset{B_1^{(m)}}{\int\cdots\int}\left[\underset{B^{(n)}}{\int\cdots\int}p(x_1,\cdots,x_n,y_1,\cdots,y_m)dx_1\cdots dx_n\right]dy_1\cdots dy_m.\end{aligned}$$

因此再由 Radon-Nikodym 定理及 (7) 知, 除去 $R^{(m)}$ 中一个 L 零测集 N 外

$$P(\xi\in B^{(n)}|\eta_1=y_1,\cdots,\eta_m=y_m)p_\eta(y_1,\cdots,y_m)$$

$$= \int_{B^{(n)}}\cdots\int p(x_1,\cdots,x_n,y_1,\cdots,y_m)dx_1\cdots dx_n, \tag{9}$$

而

$$p_\eta(y_1,\cdots,y_m) = \int_{R^{(n)}}\cdots\int p(x_1,\cdots,x_n,y_1,\cdots,y_m)dx_1\cdots dx_n,$$

则当 $(y_1,\cdots,y_m)\in N^c$ 且 $p_\eta(y_1,\cdots,y_m)\neq 0$ 时有

$$\begin{aligned}&P(\xi\in B^{(n)}|\eta_1=y_1,\cdots,\eta_m=y_m)\\&=\frac{\displaystyle\int_{B^{(n)}}\cdots\int p(x_1,\cdots,x_n,y_1,\cdots,y_m)dx_1\cdots dx_n}{\displaystyle\int_{R^{(n)}}\cdots\int p(x_1,\cdots,x_n,y_1,\cdots,y_m)dx_1\cdots dx_n},\end{aligned} \tag{10}$$

而

$$P_\eta(N\cup\{y|p_\eta(y)=0\})=0.$$

事实上,

$$\begin{aligned}P_\eta(N\cup\{y|p_\eta(y)=0\}) &\leqslant P_\eta(N)+P_\eta(\{p_\eta(y)=0\})\\&=\int_N p_\eta(y)dy+\int_{\{p_\eta(y)=0\}}p_\eta(y)dy=0,\end{aligned}$$

其中

$$y=(y_1,\cdots,y_m),\quad dy=dy_1\cdots dy_m.$$

因此在集 $N\cup\{y:p_\eta(y)=0\}$ 上 $P(\xi\in B^{(n)}|\eta_1=y_1,\cdots,\eta_m=y_m)$ 的值可以任意给定.

特别 ξ 在 $\eta=(y_1,\cdots,y_m)$ 下的条件分布函数为

$$\begin{aligned}&F(x_1,\cdots,x_n|\eta_1=y_1,\cdots,\eta_m=y_m)\\&=P(\xi_1<x_1,\cdots,\xi_n<x_n|\eta_1=y_1,\cdots,\eta_m=y_m)\\&=\frac{\displaystyle\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}p(t_1,\cdots,t_n,y_1,\cdots,y_m)dt_1,\cdots dt_n}{\displaystyle\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(t_1,\cdots,t_n,y_1,\cdots,y_m)dt_1,\cdots dt_n}.\end{aligned} \tag{11}$$

而 ξ 在 $\eta=(y_1,\cdots,y_m)$ 下的条件分布密度为

$$p(x_1,\cdots,x_n|y_1,\cdots,y_m)=\frac{p(x_1,\cdots,x_n,y_1,\cdots,y_m)}{\displaystyle\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(t_1,\cdots,t_n,y_1,\cdots,y_m)dt_1,\cdots dt_n}, \tag{12}$$

其中 $(y_1, \cdots, y_m) \in N^c$ 且, $p_\eta(y_1, \cdots, y_m) \neq 0$.

由于 $p_\eta(y_1, \cdots, y_m) = 0$ 时 (9) 的右端对任何 $B^{(n)} \in \mathscr{B}^{(n)}$ 都等于零, 因而 $p(x_1, \cdots, x_n, y_1, \cdots, y_m) = 0$, 对几乎所有的 $x \in R^{(n)}$ 都成立 (指对 L-测度). 若当 $p_\eta(y_1, \cdots, y_m) = 0$ 时令 $p(x_1, \cdots, x_n | y_1, \cdots, y_m) = p_\xi(x_1, \cdots, x_n)$, 则由 (12) 知当 $(y_1, \cdots, y_m) \in N^c$ 时, 对几乎所有的 $x \in R^{(n)}$,

$$\begin{aligned} & p(x_1, \cdots, x_n | y_1, \cdots, y_m) p_\eta(y_1, \cdots, y_m) \\ = & p(x_1, \cdots, x_n, y_1, \cdots, y_m). \end{aligned} \tag{13}$$

(注意: 实际上当 $p_\eta(y_1, \cdots, y_m) = 0$ 时, 可以令 $p(x_1, \cdots, x_n | y_1, \cdots, y_m)$ 为任一分布密度.)

例 5 设 $\zeta = (\xi_1, \cdots, \xi_n, \eta_1, \cdots, \eta_m)$ 是 $n+m$ 维连续型随机变量, 其分布密度为 $p(x_1, \cdots, x_n, y_1, \cdots, y_m)$, g 是 $(R^{(n+m)}, \mathscr{B}^{(n+m)})$ 上的可测函数且 $Eg(\zeta)$ 存在, 今求条件数学期望

$$E(g(\zeta) | \eta_1 = y_1, \cdots, \eta_m = y_m).$$

应用与例 4 相同的方法, 而将应用推论 1 的相应步骤换成应用定理 1, 即得与 (9) 相当的等式

$$\begin{aligned} & E(g(\zeta) | \eta_1 = y_1, \cdots, \eta_m = y_m) p_\eta(y_1, \cdots, y_m) \\ = & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \cdots, x_n, y_1, \cdots, y_m) p(x_1, \cdots, x_n, y_1, \cdots, y_m) dx_1 \cdots dx_n, \end{aligned}$$

其中 $(y_1, \cdots, y_m) \in (N \cup \{(y_1, \cdots, y_m) : p_\eta(y_1, \cdots, y_m) = 0\})^c$, N 为 L 零测集. 和例 4 一样可证

$$P(N \cup \{(y_1, \cdots, y_m) : p_\eta(y_1, \cdots, y_m) = 0\}) = 0.$$

由 (12) 知当 $(y_1, \cdots, y_n) \in N^c$ 且 $p_\eta(y_1, \cdots, y_m) \neq 0$ 时,

$$\begin{aligned} & E(g(\zeta) | \eta_1 = y_1, \cdots, \eta_m = y_m) \\ = & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \cdots, x_n, y_1, \cdots, y_m) p(x_1, \cdots, x_n | y_1, \cdots, y_m) dx_1 \cdots dx_n. \end{aligned} \tag{14}$$

例 6 设 (ξ, η) 为二维离散型随机变量, 且 ξ, η 分别可能取的一切值是 x_k, y_j, $k = 1, 2, \cdots$, $j = 1, 2, \cdots$, $P(\xi = x_k, \eta = y_j) = p_{kj}$, $k, j = 1, 2, \cdots$.

令 $\Omega' = \{y_j, j = 1, 2, \cdots\}$, $\mathscr{A}'$ 为 Ω' 的一切子集作成的集系, 若 $f = \eta$, P'_η 为由 η 在 $\mathscr{A}'$ 上导出的概率, 则由推论 1 知

$$P(\xi \in B | \eta = y_j) = h(y_j),$$

其中 $h(y_j)$ 满足

$$\int_{B'} h(y_j)dP'_\eta = P(\xi \in B, \eta \in B') = \sum_{\substack{x_k \in B \\ y_j \in B'}} p_{kj} = \sum_{y_j \in B'} P(\xi \in B, \eta = y_j), B' \in \mathscr{A}'.$$

而由积分定义知

$$\int_{B'} h(y_j)dP'_\eta = \sum_{y_j \in B'} h(y_j)P(\eta = y_j), B' \in \mathscr{A}'.$$

故当 $P(\eta = y_j) \neq 0$ 时

$$h(y_j) = \frac{P(\xi \in B, \eta = y_j)}{P(\eta = y_j)},$$

即

$$P(\xi \in B|\eta = y_j) = \frac{P(\xi \in B, \eta = y_j)}{P(\eta = y_j)}, P(\eta = y_j) \neq 0.$$

而当 $P(\eta = y_j) = 0$ 时,$P(\xi \in B|\eta = y_j)$ 任意给定. 因此 ξ 在 η 之下的条件分布函数

$$F(x|\eta = y_j) = P(\xi < x|\eta = y_j). \tag{15}$$

同样应用定理 1 可得

$$E(\xi|\eta = y_j) = \sum_{k=1}^{\infty} x_k P(\xi = x_k|\eta = y_j). \tag{16}$$

最后, 我们计算若干按正态分布的随机变量的条件数学期望及条件分布的公式.

例 7 设 $(\xi_1, \cdots, \xi_n)$ 是按 $N(m, D)$(D 为正定方阵) 分布的 n 维随机变量. 今往求 ξ_1 在 $\xi_2 = x_2, \cdots, \xi_n = x_n$ 下的条件分布密度及条件数学期望的公式.

由第 6 章 §6.2 知 $(\xi_2, \cdots, \xi_n)$ 是按 $N(\tilde{m}, \tilde{D})$ 分布的, 其中 $\tilde{m} = (m_2, \cdots, m_n), \tilde{D}$ 为 D 中去掉第一行及第一列作成的 $(n-1)$ 阶正定方阵. 于是由 (12) 知条件分布密度

$$p(x_1|x_2, \cdots, x_n) = \frac{|\tilde{D}|^{\frac{1}{2}}}{\sqrt{2\pi}|D|^{\frac{1}{2}}} e^{-\frac{1}{2}[(x-m)D^{-1}(x-m)' - (\tilde{x}-\tilde{m})\tilde{D}^{-1}(\tilde{x}-\tilde{m})']}, \tag{17}$$

其中 $\tilde{x} = (x_2, \cdots, x_n)$. 由于 $\tilde{D}$ 是可逆对称方阵, 故由分块矩阵的运算知: 若令方

阵 D 的第 i 行第 j 列元素为 $d_{ij}, d\hat{=}(d_{12},\cdots,d_{1n})$, 则

$$\begin{aligned}&\begin{pmatrix} 1 & -d\tilde{D}^{-1} \\ 0 & \mathrm{I} \end{pmatrix} D \begin{pmatrix} 1 & 0 \\ -\tilde{D}^{-1}d' & \mathrm{I} \end{pmatrix} \\ =&\begin{pmatrix} 1 & -d\tilde{D}^{-1} \\ 0 & \mathrm{I} \end{pmatrix}\begin{pmatrix} d_{11} & d \\ d' & \tilde{D} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -\tilde{D}^{-1}d' & \mathrm{I} \end{pmatrix} \\ =&\begin{pmatrix} d_{11}-d\tilde{D}^{-1}d' & 0 \\ 0 & \tilde{D}. \end{pmatrix}.\end{aligned} \tag{18}$$

令

$$A' = \begin{pmatrix} 1 & -d\tilde{D}^{-1} \\ 0 & \mathrm{I} \end{pmatrix},$$

则

$$A = \begin{pmatrix} 1 & 0 \\ -\tilde{D}^{-1}d' & \mathrm{I} \end{pmatrix}.$$

于是

$$D^{-1} = A\begin{pmatrix} (d_{11}-d\tilde{D}^{-1}d')^{-1} & 0 \\ 0 & \tilde{D}^{-1} \end{pmatrix} A'.$$

再由分块矩阵的运算得

$$(x-m)D^{-1}(x-m)' = \frac{1}{d_{11}-d\tilde{D}^{-1}d'}[x_1-m_1-(\tilde{x}-\tilde{m})\tilde{D}^{-1}d']^2+(\tilde{x}-\tilde{m})\tilde{D}^{-1}(\tilde{x}-\tilde{m})'.$$

易知 $|A|=1$, 故由 (18) 知 $|D|=(d_{11}-d\tilde{D}^{-1}d')|\tilde{D}|$, 由此算得 ξ_1 在 $\xi_2=x_2,\cdots,\xi_n=x_n$ 之下的条件分布密度为 $N(m_1+(\tilde{x}-\tilde{m})\tilde{D}^{-1}d', d_{11}-d\tilde{D}^{-1}d')$ 的分布密度, 因而由例 5 知

$$E(\xi_1|\xi_2=x_2,\cdots,\xi_n=x_n) = m_1+(\tilde{x}-\tilde{m})\tilde{D}^{-1}d'.$$

由数理统计学知 $R^{(n)}$ 中的超平面

$$x_1-m_1-(\tilde{x}-\tilde{m})\tilde{D}^{-1}d' = 0$$

称为 ξ_1 对 $(\xi_2,\cdots,\xi_n)$ 的回归平面 (参看 [9]§23.2).

给定函数 (或映射) 下的条件数学期望与条件概率具有 §2.2 所述的性质, 只是在叙述的形式上需要作相应的改变. 例如 §2.2 定理 1 中的 II 应叙述为: 若 $aE\xi+bE\eta$ 存在, 则 $E(a\xi+b\eta|f=\omega') = aE(\xi|f=\omega')+bE(\eta|f=\omega'), P'_f$-a.e. 其他性质我们就不再赘述了.

习题及补充

1. 试证: $\eta_T=\{\eta_t,t\in T\}$ 是概率空间 $(\Omega,\mathscr{A},P)$ 上的随机函数的充分必要条件是对每一 $t\in T,\eta_t$ 是 $(\Omega,\mathscr{A},P)$ 上的随机变量. (其中 T 是任意指标集.)

2. g 是 $(R^T,\mathscr{B}^T)$ 上的可测函数, T 是任意无穷指标集. 试证必存在 T 的可数子集 T_c, 使得对一切 $B\in\mathscr{B}^{(1)}$,

$$\{x^T:g(x^T)\in B\}\in\mathscr{B}^{T_c}\times\prod_{t\in T-T_c}R_t,$$

即函数 g 只依赖于可数个坐标

$$g(x^T)=g(x^{T_c}\times a^{T-T_c}),\text{对一切}x^T\in R^T,$$

其中 a^{T-T_c} 任意取定.

3. 设 η_T 是 $(\Omega,\mathscr{A},P)$ 上的随机函数, 若 f 是关于 $\sigma(\eta_T)$ 可测的函数, 则必存在 $\{\eta_t,t\in T\}$ 的可数子集 $\{\eta_t,t\in T_c\}\hat{=}\eta_{T_c}$, 使得 f 关于 $\sigma(\eta_{T_c})$ 可测, 因而存在 $(R^{T_c},\mathscr{B}^{T_c})$ 上的可测函数 g, 使得

$$f(\eta_T)=g(\eta_{T_c}).$$

4. 在例 3 中, 必存在 T 的可数子集 T_c 使

$$E(\xi|\eta_T)=g(\eta_{T_c})=g(\eta_{t_1},\eta_{t_2},\cdots),\text{a.e.}$$

其中 g 是 $(R^{T_c},\mathscr{B}^{T_c})$ 上的 Borel 可测函数.

5. 设 $(\Omega_i,\mathscr{A}_i,P_i),i=1,2$ 是两个概率空间, 考虑它们的乘积空间. 令 $f(\omega_1,\omega_2)=\omega_1$, 则 f 是由 $(\Omega_1\times\Omega_2,\mathscr{A}_1\times\mathscr{A}_2)$ 到 $(\Omega_1,\mathscr{A}_1)$ 的可测映射. 试证对任一 $B\in\mathscr{A}_1\times\mathscr{A}_2$,

$$E(\chi_B|f=\omega_1)=P_2(B_{\omega_1}),P_1\text{-a.e.}$$

其中 B_{ω_1} 是 B 在 ω_1 处的截集.

6. 设 $\xi=(\xi_1,\cdots,\xi_n)$ 是实随机向量, 对任意 $B\in\mathscr{B}^{(n)}$,

$$P((\xi_1,\cdots,\xi_n)\in B)=P((\xi_i,\xi_{i+1},\cdots,\xi_n,\xi_1,\cdots,\xi_{i-1})\in B),i=1,\cdots,n.$$

且对任意的 $i,j=1,\cdots,n,i\neq j$,

$$P(\xi_i=\xi_j)=0.$$

定义 $Y(x_1,\cdots,x_n)=(y_1,\cdots,y_n)=(x_i,x_{i+1},\cdots,x_n,x_1,\cdots,x_{i-1})$, 此处 $x_i=\min(x_1,\cdots,x_n)$. 则 $f(\xi_1,\cdots,\xi_n)$ 在给定函数 $Y(\xi_1,\cdots,\xi_n)=y$ 之下的条件数学期望是

$$f_0(y_1,\cdots,y_n)=\frac{1}{n}[f(y_1,\cdots,y_n)+f(y_2,\cdots,y_n,y_1)+\cdots+f(y_n,y_1,\cdots,y_{n-1})].$$

7. 如果 $E|\xi|^2 < \infty$, 我们称

$$D(\xi|\eta) \hat{=} E(|\xi - E(\xi|\eta)|^2|\eta)$$

为 ξ 在给定随机变量 η 之下的**条件方差**. 试证:

$$D\xi = E[D(\xi|\eta)] + D[E(\xi|\eta)]$$

(提示: i) 证明 $D\xi = E(E(|\xi - E\xi|^2|\eta))$; ii) 证明 $E[|\xi - E\xi|^2|\eta] = E[|\xi - E(\xi|\eta)|^2|\eta] + |E(\xi|\eta) - E\xi|^2$.)

8. 设 ξ, η 的联合分布密度为

a) $p_{\xi,\eta}(x,y) = \begin{cases} 6xy(2-x-y), & \text{若} 0 \leqslant x, y \leqslant 1, \\ 0, & \text{其他情形}. \end{cases}$

b) $p_{\xi,\eta}(x,y) = \begin{cases} \dfrac{1}{8}(y^2 - x^2)e^{-y}, & \text{若} 0 < y < \infty, |x| < y, \\ 0, & \text{其他情形}. \end{cases}$

试求: i)ξ 在给定 $\eta = y$ 之下的条件分布密度.ii)$E(\xi|\eta = x)$.

9. 设 $\xi_T = \{\xi_t, t \in T\}, T$ 是一实数集, ξ_T 是实随机函数且满足条件: 对任何 $n \geqslant 2, t_1 < t_2 < \cdots < t_n, t_k \in T, k = 1, \cdots, n$ 及实数 x

$$P(\xi_{t_n} < x|\xi_{t_1}, \cdots, \xi_{t_{n-1}}) = P(\xi_{t_n} < x|\xi_{t_{n-1}}), \text{a.e.}$$

则称此随机函数为一马氏过程. 试证对任何 $s < \tau < t$ 及 $B \in \mathscr{B}^{(1)}$ 有

$$P(\xi_t \in B|\xi_s) = E(P(\xi_t \in B|\xi_\tau)|\xi_s), \text{a.e.}$$

此方程称为Колмогоров-Chapman 方程.

10. 设 $(\xi_1, \cdots, \xi_n)$ 按 $N(m, D)$ 分布, 试证

$$\hat{E}(\xi_1|\xi_2 = x_2, \cdots, \xi_n = x_n) = m_1 + (\tilde{x} - \tilde{m})\tilde{D}^{-1}d',$$

其中 $\tilde{x}, \tilde{m}, \tilde{D}, d$ 如例 7 定义.$\hat{E}$ 如 5.2.3 末尾所定义.

§5.4　转移概率与转移测度

本节将要讨论的转移概率 (它是转移测度的特殊情形), 是研究非独立随机变量族的一个重要工具. 它与条件概率有着密切的联系. 同时, 我们注意到第 4 章所确定的乘积空间上的测度是乘积测度, 它是乘积空间上一种特殊的测度. 而一个转移

概率序列 (或有限个转移测度) 能够确定乘积空间上的非乘积测度 (Tulcea 定理), 这就推广了第 4 章的结果. 它在概率论及随机过程论中有许多应用.

定义 1 设 $(\Omega_i, \mathscr{A}_i), i=1,2$ 是可测空间, 一个映射 $\lambda: \Omega_1 \times \mathscr{A}_2 \to [0,\infty]$, 如果满足下述条件, 则称为 $(\Omega, \mathscr{A}_1)$到 $(\Omega_2, \ \mathscr{A}_2)$ **上的转移测度, 简称为** $\Omega_1 \times \mathscr{A}_2$**上的转移测度**

i) 对每个固定的 $B \in \mathscr{A}_2, \lambda(\cdot, B)$ 是 $\mathscr{A}_1$-可测函数;

ii) 对每个固定的 $\omega_1 \in \Omega_1, \lambda(\omega_1, \cdot)$ 是 $\mathscr{A}_2$ 上的测度.

如果存在 $B_n \in \mathscr{A}_2, n=1,2,\cdots$ 两两不交, $\Omega_2 = \sum\limits_{n=1}^{\infty} B_n$, 使对一切 $\omega_1 \in \Omega_1, n \geqslant 1, \lambda(\omega_1, B_n) < \infty$, 则称 λ 为 σ-**有限转移测度.**

如果对一切 $\omega_1 \in \Omega_1, \lambda(\omega_1, \Omega_2) = 1$, 则称 λ 为 $\Omega_1 \times \mathscr{A}_2$**上的转移概率.**

例 1 设 μ 是 $(\Omega_2, \mathscr{A}_2)$ 上的测度, 令

$$\lambda(\omega_1, B) = \mu(B), \omega_1 \in \Omega_1, B \in \mathscr{A}_2,$$

则 λ 是 $\Omega_1 \times \mathscr{A}_2$ 上的转移测度. 因此 $\mathscr{A}_2$ 上的测度可以看作转移测度的特殊情形. 它表示转移测度与 ω_1 无关, 而一般的转移测度表示在 $\mathscr{A}_2$ 上与 ω_1 有某种相依性的测度.

例 2 设

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \vdots & \vdots & & \vdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{pmatrix},$$

其中 $p_{ij} \geqslant 0$, 每行之和为 1, 即 $\sum\limits_j p_{ij} = 1, i=1,\cdots,k$. 设 $\Omega_1 = \Omega_2 = \{1,2,\cdots,k\}$, $\mathscr{A}_1 = \mathscr{A}_2$ 表示 $\{1,2,\cdots,k\}$ 的一切子集作成的 σ-代数, 令

$$\lambda(i, B) = \sum_{j \in B} p_{ij},$$

则 λ 是 $\Omega_1 \times \mathscr{A}_2$ 上的转移概率 (称 P 为转移概率矩阵).

注 转移概率可以如下解释: 有两个系统, 它们的一切状态分别用 Ω_1 和 Ω_2 来描述, 第二个系统的统计性质依赖于第一个系统的状态. 如果第一个系统处在状态 ω_1, 则第二个系统的状态处在集 B 中的概率用转移概率 $\lambda(\omega_1, B)$ 来描述. 例如考虑 S_1 和 S_2 两个站之间的信息传输问题, 由 S_1 发出的一切信息组成 Ω_1, S_2 收到的一切信息组成 Ω_2, 由于噪声干扰, 一个信息 $\omega_1 \in \Omega_1$, 从 S_1 发出, 在 S_2 接收到的信息未必仍是 ω_1. 我们用 $\lambda(\omega, B)$ 描述从 S_1 发出信息 ω 而在 S_2 收到信息在集 B 中的概率. 称 λ 为由 S_1 到 S_2 的信息通路的转移概率.

我们要利用转移测度建立乘积空间上的测度, 并且讨论关于这个测度的积分, 所用的方法和第 4 章中建立乘积测度及其积分的方法是一致的, 内容也有许多类似之处. 而且其特殊情形包含了第 4 章的主要结果, 请读者注意.

我们先证明

定理 1 设 f 是 $(\Omega_1\times\Omega_2,\mathscr{A}_1\times\mathscr{A}_2)$ 上的非负可测函数, λ 是 $\Omega_1\times\mathscr{A}_2$ 上的 σ-有限转移测度, 则

$$\int_{\Omega_2} f(\omega_1,\omega_2)\lambda(\omega_1,d\omega_2),\quad \omega_1\in\Omega_1, \tag{1}$$

是 $(\Omega_1,\mathscr{A}_1)$ 上的可测函数, 其中 $\lambda(\omega_1,d\omega_2)$ 表示对测度 $\lambda(\omega_1,\cdot)$ 的积分.

证 用 $\mathscr{L}$-系方法证明本定理.

由 λ 是 σ-有限转移测度知, 存在 $B_n\in\mathscr{A}_2,n=1,2,\cdots$ 两两不交,$\Omega_2=\sum\limits_{n=1}^{\infty}B_n$, 使对一切 $\omega_1\in\Omega_1,n\geqslant 1,\lambda(\omega_1,B_n)<\infty$. 欲证 $\int_{\Omega_2} f(\cdot,\omega_2)\lambda(\cdot,d\omega_2)\in\mathscr{A}_1$, 只需证对每一 n, $\int_{B_n} f(\cdot,\omega_2)\lambda(\cdot,d\omega_2)\in\mathscr{A}_1$.

令

$$\mathscr{H}=\left\{f:0\leqslant f\in\mathscr{A}_1\times\mathscr{A}_2,\text{对一切}n\geqslant 1,\int_{B_n} f(\cdot,\omega_2)\lambda(\cdot,d\omega_2)\in\mathscr{A}_1\right\},$$

则易知

i) $1\in\mathscr{H}$;

ii) $\mathscr{H}$ 对非负线性组合封闭, 且若 $f,g\in\mathscr{H}$, 有界,$f\geqslant g$, 则 $f-g\in\mathscr{H}$.

iii) 若 $\{f_n\}\in\mathscr{H},0\leqslant f_n\uparrow f$, 则 $f\in\mathscr{H}$.

iv) 设 $\mathscr{C}=\{A_1\times A_2:A_1\in\mathscr{A}_1,A_2\in\mathscr{A}_2\}$, 则 $\mathscr{C}$ 是 π-系且 $\sigma(\mathscr{C})=\mathscr{A}_1\times\mathscr{A}_2$, 而 $\mathscr{H}$ 包含 $\mathscr{C}$ 中集的示性函数.

由第 2 章 §2.2 习题 10 知 $\mathscr{H}$ 负含一切非负 $\mathscr{A}_1\times\mathscr{A}_2$ 可测函数. 定理获证. □

推论 1 设 f 是 $(\Omega_1\times\Omega_2,\mathscr{A}_1\times\mathscr{A}_2)$ 上的可测函数, λ 是 $\Omega_1\times\mathscr{A}_2$ 上的 σ-有限转移测度, 若对每一 $\omega_1\in\Omega_1$,(1) 式中的积分存在, 则 (1) 中的函数是 $(\Omega_1,\mathscr{A}_1)$ 上的可测函数.

定义 2 设 $(\Omega_i,\mathscr{A}_i),i=1,2$ 是二可测空间,λ 是 $\Omega_1\times\mathscr{A}_2$ 上的转移测度, 若存在 Ω_1 的可数分割 $\{A_m\}\subset\mathscr{A}_1,\Omega_2$ 的可数分割 $\{B_n\}\subset\mathscr{A}_2$, 使得对一切 $m,n=1,2,\cdots$ 有

$$\sup_{\omega_1\in A_m}\lambda(\omega_1,B_n)<\infty,$$

则称 λ 为一致 σ-有限转移测度.

定理 2 设 $(\Omega_i,\mathscr{A}_i), i=1,2,\cdots,n$ 是 n 个可测空间, $\Omega^{(k)}=\prod_{i=1}^{k}\Omega_i, \mathscr{A}^{(k)}=\prod_{i=1}^{k}\mathscr{A}_i, k=1,\cdots,n$. 若 λ_1 是 $\mathscr{A}_1$ 上的 σ-有限测度,λ_k 是 $\Omega^{(k-1)}\times\mathscr{A}_k$ 上的 σ-有限转移测度,$k=2,\cdots,n$. 则

$$\lambda^{(n)}(B)=\int_{\Omega_1}\cdots\int_{\Omega_n}\chi_B(\omega_1,\cdots,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)\cdots\lambda_1(d\omega_1), B\in\mathscr{A}^{(n)} \tag{2}$$

是 $\mathscr{A}^{(n)}$ 上的测度. 若 $\lambda_2,\cdots,\lambda_n$ 是一致 σ-有限的, 则 $\lambda^{(n)}$ σ-有限.

证 由于 $\mathscr{A}^{(k)}=\mathscr{A}^{(k-1)}\times\mathscr{A}_k, k=2,\cdots,n, \chi_B$ 是 $\mathscr{A}^{(n)}$ 可测的, 反复应用定理 1 知 $\lambda^{(n)}(B)$ 有意义, 因而是 $\mathscr{A}^{(n)}$ 上的非负集函数.

再由 (2) 式左端最里层积分开始, 依次对每一重积分应用单调收敛定理, 即知 $\lambda^{(n)}$ 是 σ-可加的.

对于 $\lambda^{(n)}$ 的 σ-有限性, 我们只证 $n=2$ 的情形, 一般情形留作习题.

由于 λ_1 是 σ-有限的, 故存在 Ω_1 的可数分割 $\{A_n\}\subset\mathscr{A}_1$, 使 $\lambda_1(A_n)<\infty$, 再由 λ_2 是一致 σ-有限的转移测度, 因而存在 Ω_1 的可数分割 $\{B_m\}\subset\mathscr{A}_1$ 及 Ω_2 的可数分割 $\{C_k\}\subset\mathscr{A}_2$, 使 $\sup\limits_{\omega_1\in B_m}\lambda_2(\omega_1,C_k)<\infty$, 则

$$(A_n\cap B_m)\times C_k\in\mathscr{A}_1\times\mathscr{A}_2, n,m,k=1,2,\cdots$$

是 $\Omega_1\times\Omega_2$ 的可数分割, 且对每一 n,m,k

$$\begin{aligned}\int_{\Omega_1}\int_{\Omega_2}\chi_{(A_n\cap B_m)\times C_k}(\omega_1,\omega_2)\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)&=\int_{\Omega_1}\chi_{A_n\cap B_m}(\omega_1)\lambda_2(\omega_1,C_k)\lambda_1(d\omega_1)\\&\leqslant\sup_{\omega_1\in B_m}\lambda_2(\omega_1,C_k)\cdot\lambda_1(A_n)<\infty.\end{aligned}$$

因而 $\lambda^{(2)}(A_n\cap B_m\times C_k)<\infty$, 即 $\lambda^{(2)}$ 是 $\mathscr{A}_1\times\mathscr{A}_2$ 上的 σ-有限测度 □

显然, 定理 2 是乘积测度存在定理的推广, 且若 λ_1 是概率,$\lambda_2,\cdots,\lambda_n$ 是转移概率, 则由 (2) 定义的 $\lambda^{(n)}$ 是 $(\Omega^{(n)},\mathscr{A}^{(n)})$ 上的概率.

下一定理的作用类似于重积分化累次积分.

定理 3 设 $\lambda^{(n)}$ 是由定理 2 所确定的 $\mathscr{A}^{(n)}$ 上的测度, f 是 $(\Omega^{(n)},\mathscr{A}^{(n)})$ 上的可测函数, 则 f 关于 $\lambda^{(n)}$ 积分存在的充分必要条件是

$$\int_{\Omega_1}\cdots\int_{\Omega_n}f^{\pm}(\omega_1,\cdots,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)\cdots\lambda_1(d\omega_1)$$

有一有限. 且当积分存在时

$$\begin{aligned}&\int_{\Omega^{(n)}} f(\omega_1,\cdots,\omega_n)d\lambda^{(n)}\\&=\int_{\Omega_1}\cdots\int_{\Omega_n} f(\omega_1,\cdots,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)\cdots\lambda_1(d\omega_1)\end{aligned}\tag{3}$$

证　当 $f=\chi_B, B\in\mathscr{A}^{(n)}$ 时 (3) 即 (2). 由积分线性性知 f 是非负简单函数时 (3) 成立, 再由单调收敛定理知 (3) 对非负 $\mathscr{A}^{(n)}$-可测函数成立. 当 f 为实可测函数时,$f=f^+-f^-$, 由于 f 关于 $\lambda^{(n)}$ 积分存在的充分必要条件是$\int_{\Omega^{(n)}} f^{\pm}d\lambda^{(n)}$ 有一有限. 因而为证 (3) 对实可测函数 f 成立, 只需证下列命题：对一切自然数 n, 若 f,f_1,f_2 是 $(\Omega^{(n)},\mathscr{A}^{(n)})$ 上的可测函数且 $f=f_1-f_2,\lambda^{(n)}$-a.e., $f_i\geqslant 0, i=1,2$, $\int_{\Omega^{(n)}} f_i d\lambda^{(n)}$ 有一有限, 则 (3) 成立 $\left(\text{为确定起见, 设}\int_{\Omega^{(n)}} f_2 d\lambda^{(n)}<\infty\right)$.

显然 $n=1$ 时上述命题成立. 设上述命题对 $n-1$ 成立, 令

$$f_i'(\omega_1,\cdots,\omega_{n-1})=\int_{\Omega_n} f_i(\omega_1,\cdots,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},d\omega_n), i=1,2.$$

由于 (3) 对非负可测函数成立, 故有

$$\int_{\Omega^{(n-1)}} f_2' d\lambda^{(n-1)}=\int_{\Omega^{(n)}} f_2 d\lambda^{(n)}<\infty.\tag{4}$$

因此 $f_2', \lambda^{(n-1)}$-a.e. 有限, 于是由积分的线性性

$$f'(\omega_1,\cdots,\omega_{n-1})=\int_{\Omega_n} f(\omega_1,\cdots,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)$$

$\lambda^{(n-1)}$-a.e. 有意义且由本节习题 2 可证

$$f'=f_1'-f_2', \lambda^{(n-1)}\text{-a.e.}\tag{5}$$

于是由 (5) 及归纳假设有

$$\int_{\Omega^{n-1}} f' d\lambda^{(n-1)}=\int_{\Omega_1}\cdots\int_{\Omega_{n-1}} f'\lambda_{n-1}(\omega_1,\cdots,\omega_{n-2},d\omega_{n-1})\cdots\lambda_1(d\omega_1).\tag{6}$$

由 (4), (5), 积分的线性性及 (3) 对非负函数成立

$$\begin{aligned}\int_{\Omega^{(n-1)}} f' d\lambda^{(n-1)}&=\int_{\Omega^{(n-1)}} f_1' d\lambda^{(n-1)}-\int_{\Omega^{n-1}} f_2' d\lambda^{(n-1)}\\&=\int_{\Omega_1}\cdots\int_{\Omega_n} f_1(\omega_1,\cdots,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)\cdots\lambda_1(d\omega_1)\end{aligned}$$

$$-\int_{\Omega_1}\cdots\int_{\Omega_n} f_2(\omega_1,\cdots,\omega_n)\lambda_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)\cdots\lambda_1(d\omega_1)$$
$$=\int_{\Omega^{(n)}} f_1 d\lambda^{(n)}-\int_{\Omega^{(n)}} f_2 d\lambda^{(n)}$$
$$=\int_{\Omega^{(n)}} f d\lambda^{(n)}. \tag{7}$$

最后由 (6),(7) 及 f' 的定义即知 (3) 成立. 定理获证. □

定理 4 若 $(\Omega_i,\mathscr{A}_i), i=1,2$ 为二可测空间,λ_1 是 $\mathscr{A}_1$ 上的 σ-有限测度, λ_2 是 $\Omega_1\times\mathscr{A}_2$ 上的 σ-有限转移测度, $\lambda^{(2)}$ 为由 λ_1,λ_2 按定理 2 决定的测度, f 是 $(\Omega_2,\mathscr{A}_2)$ 上的可测函数且 $\int_{\Omega^{(2)}} f d\lambda^{(2)}$ 存在 (将 f 看作 $\Omega_1\times\Omega_2$ 上的函数), 若令

$$\mu(A)=\int_{\Omega_1}\lambda_2(\omega_1,A)\lambda_1(d\omega_1), A\in\mathscr{A}_2, \tag{8}$$

则 μ 是 $\mathscr{A}_2$ 上的测度且

$$\int_{\Omega_1}\int_{\Omega_2} f(\omega_2)\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)=\int_{\Omega_2} f(\omega_2)\left[\int_{\Omega_1}\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)\right]. \tag{9}$$

证 由单调收敛定理及转移测度的定义易证 μ 是 $\mathscr{A}_2$ 上的测度.

若 $f(\omega_2)=\chi_A(\omega_2), A\in\mathscr{A}_2$, 则 (9) 式左右两端皆等于 $\mu(A)$, 故 (9) 式对 $\mathscr{A}_2$-可测的示性函数成立, 由积分可加性及单调收敛定理易证 (9) 式当 f 是非负可测时也成立.

若 f 是 $\mathscr{A}_2$ 实可测函数, $f=f^+-f^-$, 由 $\int_{\Omega^{(2)}} f d\lambda^{(2)}$ 存在知 $\int_{\Omega^{(2)}} f^{\pm}(\omega_2)d\lambda^{(2)}$ 有一有限, 不妨设 $\int_{\Omega^{(2)}} f^- d\lambda^{(2)}$ 有限. 由定理 3 及 (9) 对非负可测函数成立知

$$\int_{\Omega_2} f^-(\omega_2)d\mu=\int_{\Omega_2} f^-(\omega_2)\left[\int_{\Omega_1}\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)\right]$$
$$=\int_{\Omega_1}\int_{\Omega_2} f^-(\omega_2)\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)=\int_{\Omega^{(2)}} f^- d\lambda^{(2)}<\infty,$$

因而 (9) 式两端的积分皆存在. 再由定理 3、积分的线性性及 (9) 式对非负可测函数成立得

$$\int_{\Omega_1}\int_{\Omega_2} f(\omega_2)\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)$$
$$=\int_{\Omega^{(2)}} f^+ d\lambda^{(2)}-\int_{\Omega^{(2)}} f^- d\lambda^{(2)}$$

$$=\int_{\Omega_1}\int_{\Omega_2} f^+(\omega_2)\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)-\int_{\Omega_1}\int_{\Omega_2} f^-(\omega_2)\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1)$$

$$=\int_{\Omega_2} f^+(\omega_2)d\mu-\int_{\Omega_2} f^-(\omega_2)d\mu=\int_{\Omega_2} fd\mu$$

$$=\int_{\Omega_2} f(\omega_2)\left[\int_{\Omega_1}\lambda_2(\omega_1,d\omega_2,)\lambda_1(d\omega_1)\right].$$

对于 f 是实函数的情形 (9) 式获证. 对 f 是复函数的情形按照通常的方法同样可证 (9) 式成立. □

现在证明著名的 Tulcea 定理, 它是无穷乘积概率定理的推广, 其证明方法也几乎与无穷乘积概率定理的证明相同.

定理 5(Tulcea 定理) 设 $(\Omega_n,\mathscr{A}_n), n=1,2,\cdots$ 是可测空间, $\Omega^{(n)}=\prod_{k=1}^{n}\Omega_k$, $\mathscr{A}^{(n)}=\prod_{k=1}^{n}\mathscr{A}_k, \Omega^{(\infty)}=\prod_{k=1}^{\infty}\Omega_k, \mathscr{A}^{(\infty)}=\prod_{k=1}^{\infty}\mathscr{A}_k. P_n(\omega_1,\cdots,\omega_{n-1},A_n),(\omega_1,\cdots,\omega_{n-1},A_n)\in\Omega^{(n-1)}\times\mathscr{A}_n$ 是 $\Omega^{(n-1)}\times\mathscr{A}_n$ 上的转移概率, $n=2,3,\cdots,P_1(A),A\in\mathscr{A}_1$ 是 $\mathscr{A}_1$ 上的概率, 则在 $\mathscr{A}^{(\infty)}$ 上有唯一的概率 $P^{(\infty)}$ 存在, 使得

$$P^{(\infty)}(C(B^{(n)}))=P^{(n)}(B^{(n)}), \tag{10}$$

其中 $C(B^{(n)})$ 表示 $\mathscr{A}^{(\infty)}$ 中以 $B^{(n)}$ 为底的柱集, $B^{(n)}\in\mathscr{A}^{(n)}$ 而

$$P^{(n)}(B^{(n)})=\int_{\Omega_1}\cdots\int_{\Omega_n}\chi_{B^{(n)}}(\omega_1,\cdots,\omega_n)P_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)\cdots P_1(d\omega_1). \tag{11}$$

证 首先我们注意由一切 $C(B^{(n)}),B^{(n)}\in\mathscr{A}^{(n)},n$ 为任意正整数, 做成的类为一集代数, 记此类为 $\mathscr{D}$, 则 $\mathscr{A}^{(\infty)}=\sigma(\mathscr{D})$. 因此欲证 $P^{(\infty)}$ 的存在性及唯一性, 只需证明

I) (10) 在 $\mathscr{D}$ 上定义了一个有限可加集函数 $P^{(\infty)}$;

II) $P^{(\infty)}$ 在 $\mathscr{D}$ 上在 $\varnothing$ 处连续.

因为若 I)II) 获证, 则 $P^{(\infty)}$ 是 $\mathscr{D}$ 上的一个 σ-可加集函数, 且由 (10) 知显然有 $P^{(\infty)}$ 非负及 $P^{(\infty)}(\Omega^{(\infty)})=1$, 因此 $P^{(\infty)}$ 是 $\mathscr{D}$ 上的概率. 于是由测度扩张定理知 $P^{(\infty)}$ 在 $\mathscr{A}^{(\infty)}$ 上是唯一满足 (10) 的概率.

现在我们来证明 I 和 II.

i) 首先证明 (10) 不因 $C(B^{(n)})$ 的表示法不同而不同. 设 $C(B^{(n)})=C(B^{(m)})$, 且 $m<n$, 则必有

$$B^{(n)}=B^{(m)}\times\Omega_{m+1}\times\cdots\times\Omega_n,$$

于是由

$$\int_{\Omega_k}\chi_{B^{(m)}\times\Omega_{m+1}\times\cdots\times\Omega_k}(\omega_1,\cdots,\omega_k)P_k(\omega_1,\cdots,\omega_{k-1}d\omega_k)$$
$$=\chi_{B^{(m)}\times\Omega_{m+1}\times\cdots\times\Omega_{k-1}}(\omega_1,\cdots,\omega_{k-1}),m+1\leqslant k\leqslant n,$$

及 (11) 知

$$P^{(n)}(B^{(n)})=P^{(n)}(B^{(m)}\times\Omega_{m+1}\times\cdots\times\Omega_n)$$
$$=P^{(n-1)}(B^{(m)}\times\Omega_{m+1}\times\cdots\times\Omega_{n-1})=\cdots=P^{(m)}(B^{(m)}).$$

因此 (10) 在 $\mathscr{D}$ 上定义了集函数 $P^{(\infty)}$. 容易证明 $P^{(\infty)}$ 在 $\mathscr{D}$ 上是有限可加的. 因为若 $A,B\in\mathscr{D},A\cap B=\varnothing$, 则 $A=B_1^{(m)}\times\prod\limits_{k=m+1}^{\infty}\Omega_k,B=B^{(n)}\times\prod\limits_{k=n+1}^{\infty}\Omega_k$, 且不妨设 $m\leqslant n$, 令 $A^{(n)}=B_1^{(m)}\times\Omega_{m+1}\times\cdots\times\Omega_n$, 于是 $A^{(n)}\cap B^{(n)}=\varnothing$,

$$P^{(\infty)}(A+B)=P^{(n)}(A^{(n)}+B^{(n)})=P^{(n)}(A^{(n)})+P^{(n)}(B^{(n)})$$
$$=P^{(m)}(B_1^{(m)})+P^{(n)}(B^{(n)})=P^{(\infty)}(A)+P^{(\infty)}(B).$$

因此 I 获证.

ii) 现在往证 II, 先作两点注释.

第一, 反复应用定理 1 知

$$\int_{\Omega_{k+1}}\cdots\int_{\Omega_n}\chi_{B^{(n)}}(\omega_1,\cdots,\omega_n)P_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)\cdots P_{k+1}(\omega_1,\cdots,\omega_k,d\omega_{k+1})\tag{12}$$

是 $\mathscr{A}^{(k)}$ 可测函数;

第二, 用定理 2 完全同样的证明方法可以证明对任何给定的 $(\omega_1,\cdots,\omega_k)\in\Omega^{(k)}$,(12) 定义了 $B^{(n)}(\omega_1,\cdots,\omega_k)$ (即 $B^{(n)}$ 在 $(\omega_1,\cdots,\omega_k)$ 处的截集) 的类上的概率, 即定义了 $\mathscr{A}_{k+1}\times\cdots\times\mathscr{A}_n$ 上的一个概率. 因此对任意给定的 k 及 $(\omega_1,\cdots,\omega_k)\in\Omega^{(k)}$ 来说, 等式

$$Q^{(k)}(C(B^{(n)})(\omega_1,\cdots,\omega_k))$$
$$=\begin{cases}\displaystyle\int_{\Omega_{k+1}}\cdots\int_{\Omega_n}\chi_{B^{(n)}}(\omega_1,\cdots,\omega_n)P_n(\omega_1,\cdots,\omega_{n-1},d\omega_n)\cdots P_{k+1}(\omega_1,\cdots,\omega_k,\\ \qquad d\omega_{k+1}),\text{若}k<n,\\ \chi_{B^{(n)}\times\Omega_{n+1}\times\cdots\times\Omega_k}(\omega_1,\cdots,\omega_k),\text{若}k\geqslant n.\end{cases}\tag{13}$$

定义了 $\prod\limits_{j=k+1}^{\infty}\mathscr{A}_j$ 中可测柱集类上的一个有限可加概率, 且 $Q^{(k)}(C(B^{(n)})(\omega_1,\cdots,\omega_k))$ 是一 $\mathscr{A}^{(k)}$-可测函数.

要想证明 II, 只需证明: 若 $A_j \in \mathscr{D}, A_j \downarrow$, 且存在 $\varepsilon > 0$ 使 $P^{(\infty)}(A_j) \geqslant \varepsilon(j = 1,2,\cdots)$, 则 $\bigcap_{j=1}^{\infty} A_j \neq \varnothing$.

首先令

$$B_j^{(1)} = \left\{\omega_1 : \omega_1 \in \Omega_1, Q^{(1)}(A_j(\omega_1)) \geqslant \frac{\varepsilon}{2}\right\},$$

则由上述注释知 $Q^{(1)}(A_j(\omega_1))$ 是 $\mathscr{A}_1$-可测函数, 因而 $B_j^{(1)} \in \mathscr{A}_1$, 故由 (10), (11), (13) 知

$$\begin{aligned}
\varepsilon &\leqslant P^{(\infty)}(A_j) = \int_{\Omega_1} Q^{(1)}(A_j(\omega_1))P_1(d\omega_1) \\
&= \int_{B_j^{(1)}} Q^{(1)}(A_j(\omega_1))P_1(d\omega_1) + \int_{[B_j^{(1)}]^c} Q^{(1)}(A_j(\omega_1))P_1(d\omega_1) \\
&\leqslant P_1(B_j^{(1)}) + \int_{\Omega_1} \frac{\varepsilon}{2} P_1(d\omega_1) = P_1(B_j^{(1)}) + \frac{\varepsilon}{2}, \\
&\quad j = 1,2,\cdots,
\end{aligned}$$

因而 $P_1(B_j^{(1)}) \geqslant \frac{\varepsilon}{2}, j = 1,2,\cdots$.

再由 $A_j(\omega_1) \downarrow$ 知 $B_j^{(1)} \downarrow$, 故由 P_1 是 $\mathscr{A}_1$ 上的概率知 $\bigcap_{j=1}^{\infty} B_j^{(1)} \neq \varnothing$, 即有一 $\bar{\omega}_1 \in B_j^{(1)}, j = 1,2,\cdots$, 亦即存在 $\bar{\omega}_1 \in \Omega_1$ 使

$$Q^{(1)}(A_j(\bar{\omega}_1)) \geqslant \frac{\varepsilon}{2}, j = 1,2,\cdots, A_j(\bar{\omega}_1) \downarrow. \tag{14}$$

由 (14) 并注意到 $(A_j(\bar{\omega}_1))(\omega_2) = A_j(\bar{\omega}_1, \omega_2)$,

$$Q^{(1)}(A_j(\bar{\omega}_1)) = \int_{\Omega_2} Q^{(2)}(A_j(\bar{\omega}_1, \omega_2))P_2(\bar{\omega}_1, d\omega_2),$$

重复上述论证即得有一 $\bar{\omega}_2 \in \Omega_2$ 存在, 使

$$Q^{(2)}(A_j(\bar{\omega}_1, \bar{\omega}_2)) \geqslant \frac{\varepsilon}{2^2}, \quad j = 1,2,\cdots, A_j(\bar{\omega}_1, \bar{\omega}_2) \downarrow.$$

应用归纳法可证有 $(\bar{\omega}_1, \bar{\omega}_2, \cdots) \in \Omega^{(\infty)}$ 存在, 使得对任何正整数 N 来说

$$Q^{(N)}(A_j(\bar{\omega}_1, \cdots, \bar{\omega}_N)) \geqslant \frac{\varepsilon}{2^N}, \quad j = 1,2,\cdots.$$

设 $A_j = C(B^{(N_j)}), j = 1,2,\cdots$, 故由上式知 $(\bar{\omega}_1, \cdots, \bar{\omega}_{N_j}) \in B^{(N_j)}, j = 1,2,\cdots$, 因而对每一 j 来说, $(\bar{\omega}_1, \bar{\omega}_2, \cdots) \in A_j$, 即 $\bigcap_{j=1}^{\infty} A_j \neq \varnothing$. 定理获证. □.

习题及补充

1. 设 $f(x,y)$ 是 $[0,1]\times[0,1]$ 上的非负连续函数, 则对任意 Borel 集 $B\subset[0,1]$, 令

$$\lambda(x,B)=\int_B f(x,y)dy,$$

其中 dy 表示对 Lebesgue 测度的积分. 则 λ 是 $[0,1]\times([0,1]\cap\mathscr{B}^{(1)})$ 上的转移测度.

2. 设 $(\Omega_i,\mathscr{A}_i), i=1,2$ 是可测空间, λ_1 是 $\mathscr{A}_1$ 上的 σ-有限测度, λ_2 是 $\Omega_1\times\mathscr{A}_2$ 上的 σ-有限转移测度,

$$\nu(B)=\int_{\Omega_1}\int_{\Omega_2}\chi_B(\omega_1,\omega_2)\lambda_2(\omega_1,d\omega_2)\lambda_1(d\omega_1), B\in\mathscr{A}_1\times\mathscr{A}_2,$$

f 是 $\mathscr{A}_1\times\mathscr{A}_2$ 可测函数, 则 $f=0,\nu$-a.e. 的充分必要条件是存在一个 λ_1 零测集 N, 使得对一切 $\omega_1\in N^c$,

$$f(\omega_1,\cdot)=0, \lambda_2(\omega_1,\cdot)\text{-a.e.}$$

3. 设 $(\Omega_i,\mathscr{A}_i)$ 是可测空间,$i=1,2,3.\lambda$ 是 $\Omega_2\times\mathscr{A}_3$ 上的 σ-有限转移测度, f 是 $\mathscr{A}_1\times\mathscr{A}_3$ 可测函数, 若积分

$$g(\omega_1,\omega_2)\hat{=}\int_{\Omega_3}f(\omega_1,\omega_3)\lambda(\omega_2,d\omega_3), (\omega_1,\omega_2)\in\Omega_1\times\Omega_2$$

对一切 ω_1,ω_2 存在, 则 g 是 $\mathscr{A}_1\times\mathscr{A}_2$ 可测的.

4. 设 $(\Omega_i,\mathscr{A}_i), i=1,2,3,4$ 是可测空间, λ_1 和 λ_2 分别是 $\Omega_1\times\mathscr{A}_2, \Omega_2\times\mathscr{A}_3$ 上的转移概率, 则

$$\lambda_1\circ\lambda_2(\omega_1,B)\hat{=}\int\lambda_2(\omega_2,B)\lambda_1(\omega_1,d\omega_2), (\omega_1,B)\in\Omega_1\times\mathscr{A}_3$$

是 $\Omega_1\times\mathscr{A}_3$ 上的转移概率. 如果 λ_3 是 $\Omega_3\times\mathscr{A}_4$ 上的转移概率, 则 $(\lambda_1\circ\lambda_2)\circ\lambda_3=\lambda_1\circ(\lambda_2\circ\lambda_3)$.

5. 设 $\Omega=\{1,\cdots,n\},\mathscr{A}$ 是 Ω 的一切子集作成的类, 令 λ_1 和 λ_2 都是 $\Omega\times\mathscr{A}$ 上的转移概率, 它们分别由转移概率矩阵 P_1,P_2 所决定. 则与 $\lambda_1\circ\lambda_2$ 相应的转移概率矩阵是 $P_1\cdot P_2$(参看例 2).

6. 设 $(\Omega,\mathscr{A})$ 为可测空间,$\mathscr{C}$ 为 $\mathscr{A}$ 的子 σ-代数, 称 $\pi(x,A)$ 为一概率核, 若对每一 $A\in\mathscr{A}$ 为 $\mathscr{C}$ 可测函数, 对每一 $x\in\Omega$ 是 $\mathscr{A}$ 上的概率. 设 ν 是 $\mathscr{A}$ 上任一概率, 试证

$$\nu_\pi(A)\hat{=}\int\pi(x,A)\nu(dx)$$

为 $(\Omega,\mathscr{A})$ 上的概率测度. 它可以看作由 $\mathscr{C}$ 上的概率到 $\mathscr{A}$ 上的扩张, 也可以看作一种预报.

7. 若 μ 是 $(\Omega,\mathscr{A})$ 上的概率, $\mathscr{C}$ 是 $\mathscr{A}$ 的子 σ-代数,$\pi(x,F),(x,F)\in\Omega\times\mathscr{A}$ 是关于 $\mathscr{C}$ 可测的概率核 (即 $(\Omega,\mathscr{C})$ 到 $(\Omega,\mathscr{A})$ 上的转移概率.), 试证对 $F\in\mathscr{A}$,

$$E_\mu(\chi_F|\mathscr{C})=\pi(\cdot,F)\mu\text{-a.e.} \tag{15}$$

与

$$\mu(F\cap G)=\int_G\pi(x,F)\mu(dx),G\in\mathscr{C} \tag{16}$$

等价.

8. 证明定理 2 中 $\lambda^{(n)}$ 的 σ-有限性.

§5.5 正则条件概率、条件分布及 Колмогоров 和谐定理

我们注意到 §5.2 所定义的条件概率与条件数学期望具有与概率、数学期望非常类似的性质. 即

1° 函数 $P(A|\mathscr{C})$ 具有性质

$$P(\Omega|\mathscr{C})=1,\text{a.e.},P(A|\mathscr{C})\geqslant 0,\text{a.e.}$$

$$P(\sum A_k|\mathscr{C})=\sum P(A_k|\mathscr{C}),\text{a.e.}$$

其中 $A_k,k=1,2,\cdots$ 两两不交.

2° 函数 $E(\xi|\mathscr{C})$a.e. 具有 ξ 关于 $P(\xi|\mathscr{C})$ 的积分的性质, 即

$$E\left(\sum_{k=1}^n a_k\chi_{A_k}|\mathscr{C}\right)=\sum_{k=1}^n a_kP(A_k|\mathscr{C}),\text{a.e.}\quad 0\leqslant\xi_n\uparrow\xi\Rightarrow E(\xi_n|\mathscr{C})\uparrow E(\xi|\mathscr{C}),\text{a.e.}$$

但是并不能由 1°,2° 直接得出 $E(\xi|\mathscr{C})$ 是 ξ 关于 $P(A|\mathscr{C}),A\in\mathscr{A}$, 的积分. 因为 $E(\xi|\mathscr{C})$ 及 $P(A|\mathscr{C})$ 都是 ω 的函数, 所以谈到 $E(\xi|\mathscr{C})$ 是 ξ 关于 $P(A|\mathscr{C}),A\in\mathscr{A}$, 的积分的含意应该是指对每一 $\omega\in\Omega$, 或至少存在一个与 A 无关的零概率集 N, 使得对每一 $\omega\in N^c,E(\xi|\mathscr{C})(\omega)$ 是 ξ 关于 $P(\cdot|\mathscr{C})(\omega)$ 的积分. 这就首先要求对每个 $\omega\in N^c,P(A|\mathscr{C})(\omega),A\in\mathscr{A}$ 是 $\mathscr{A}$ 上的概率. 但是由 §5.2 定理 1′ 后面的注释知 1° 中诸性质中的例外集一般来说是与给定的集 $A,A_k,k=1,2,\cdots$ 有关的. 而 $\mathscr{A}$ 中元未必只有可数多个, 因此未必存在共同的例外零概率集.

为了解决上述问题, 我们将在下面引进正则条件概率及条件分布的概念. 并进一步证明当 $P(A|\mathscr{C})(\omega)$ 是正则条件概率时, $E(\xi|\mathscr{C})(\omega)$ 是 ξ 关于 $P(\cdot|\mathscr{C})(\omega)$ 的数学期望. 最后还证明在一定意义下正则条件概率及条件分布是永远存在的. 在证明后一问题的过程中, 还将证明在随机过程论中具有基础意义的 Колмогоров 和谐定理.

5.5.1 正则条件概率、条件分布及混合条件分布的定义及性质

一、正则条件概率: 设 $(\Omega,\mathscr{A},P)$ 是概率空间,$\mathscr{C}$ 是 $\mathscr{A}$ 的子 σ-代数.

定义 1 设 $P^{\mathscr{C}}(\omega,A)$ 是定义在 $\Omega\times\mathscr{A}$ 上的函数, 满足

i) 对每一 $A\in\mathscr{A},P^{\mathscr{C}}(\cdot,A)$ 是 $\mathscr{C}$-可测函数, 对每一 $\omega\in\Omega,P^{\mathscr{C}}(\omega,\cdot)$ 是 $\mathscr{A}$ 上的概率;

ii) 对每一 $A\in\mathscr{A},P^{\mathscr{C}}(\cdot,A)$ 是 A 在 $\mathscr{C}$ 之下 (关于 P) 的条件概率, 即对一切 $B\in\mathscr{C}$

$$\int_B P^{\mathscr{C}}(\omega,A)dP=P(A\cap B).$$

则称 $P^{\mathscr{C}}(\omega,A),(\omega,A)\in\Omega\times\mathscr{A}$ 为在 $\mathscr{C}$ 之下 (关于 P) 的正则条件概率.

注意: 对可测空间 $(\Omega,\mathscr{C})$ 和 $(\Omega,\mathscr{A})$, 正则条件概率是 $\Omega\times\mathscr{A}$ 上的转移概率.

当 $P^{\mathscr{C}}(\omega,A)$ 是正则条件概率时, 引言中所提出的问题回答是肯定的. 即

定理 1 若 $P^{\mathscr{C}}(\omega,A)$ 是在 $\mathscr{C}$ 之下关于 P 的正则条件概率, 且 ξ 是使 $E\xi$ 存在的随机变量, 则

$$E(\xi|\mathscr{C})(\omega)=\int_\Omega \xi(\omega')P^{\mathscr{C}}(\omega,d\omega'),\quad \omega\in N^c, \tag{1}$$

其中 $P(N)=0,N\in\mathscr{C},N$ 与 ξ 有关, 而 $\int_\Omega \xi(\omega')P^{\mathscr{C}}(\omega,d\omega')$ 表示给定 ω 时 ξ 对概率 $P^{\mathscr{C}}(\omega,\cdot)$ 的积分.

证 首先, 若 $\xi=\chi_A,A\in\mathscr{A}$, 则

$$E(\xi|\mathscr{C})=E(\chi_A|\mathscr{C})=P(A|\mathscr{C})=P^{\mathscr{C}}(\cdot,A),\text{a.e.}$$

故存在 $N\in\mathscr{C},P(N)=0$, 当 $\omega\in N^c$ 时

$$\begin{aligned}E(\xi|\mathscr{C})(\omega)&=P^{\mathscr{C}}(\omega,A)=\int_\Omega \chi_A(\omega')P^{\mathscr{C}}(\omega,d\omega')\\&=\int_\Omega \xi(\omega')P^{\mathscr{C}}(\omega,d\omega').\end{aligned}$$

即 (1) 对 ξ 为 $\mathscr{A}$ 中集的示性函数时成立.

其次, 由条件期望的线性性质及积分的线性性质知当 ξ 是非负简单随机变量时 (1) 式成立. 再由条件期望的单调收敛性及积分的单调收敛定理知当 ξ 是非负随机变量时 (1) 式成立.

再次若 ξ 是 $E\xi$ 存在的实随机变量,$\xi=\xi^+-\xi^-,E\xi^-<\infty$, 则

$$\begin{aligned}E(\xi^+|\mathscr{C})(\omega) &= \int_\Omega \xi^+(\omega')P^{\mathscr{C}}(\omega,d\omega'),\omega\in N_1^c,\\ E(\xi^-|\mathscr{C})(\omega) &= \int_\Omega \xi^-(\omega')P^{\mathscr{C}}(\omega,d\omega')<\infty,\omega\in N_2^c,\\ E(\xi|\mathscr{C})(\omega) &= E(\xi^+|\mathscr{C})(\omega)-E(\xi^-|\mathscr{C})(\omega),\omega\in N_3^c.\end{aligned}$$

于是

$$E(\xi|\mathscr{C})(\omega)=\int_\Omega \xi(\omega)P^{\mathscr{C}}(\omega,d\omega'),\omega\in(N_1\cup N_2\cup N_3)^c,$$

其中 $N_i\in\mathscr{C},P(N_i)=0,i=1,2,3$. 令 $N=N_1\cup N_2\cup N_3$. 则 $P(N)=0,N\in\mathscr{C}$. 因而当 ξ 是实随机变量时 (1) 成立.

最后当 ξ 为复随机变量且 $E\xi$ 存在时, 分实虚部仍可证 (1) 式成立. □

注　此定理去掉 ξ 有限的要求仍然成立.

由 5.2.3 的定理 7 立得

定理 2　设 $\mathscr{C}\subset\mathscr{C}'\subset\mathscr{A},P^{\mathscr{C}}(\omega,A),P^{\mathscr{C}'}(\omega,A),(\omega,A)\in\Omega\times\mathscr{A}$ 分别是在 $\mathscr{C}$ 和 $\mathscr{C}'$ 之下关于 P 的正则条件概率. 设 ξ 是 $\mathscr{A}$ 可测的, ξ' 是 $\mathscr{C}'$ 可测的. 若 $E\xi\xi'$ 及 $E\xi$ 存在 (若 ξ,ξ' 皆为复可测函数时要求 $\xi\xi',\xi$ 可积), 则存在 $N\in\mathscr{C},P(N)=0$, 对任一 $\omega\in N^c$

$$\int_\Omega \xi'(\omega')\xi(\omega')P^{\mathscr{C}}(\omega,d\omega')=\int_\Omega\left[\int\xi(\omega')P^{\mathscr{C}'}(\omega'',d\omega')\right]\xi'(\omega'')P^{\mathscr{C}}(\omega,d\omega''). \tag{2}$$

证　由 5.2.3 定理 7、定理 8 及本节定理 1 知 (2) 式左端 a.e. 等于

$$\begin{aligned}E(\xi\xi'|\mathscr{C})(\cdot) &= E(E(\xi\xi'|\mathscr{C}')|\mathscr{C})(\cdot)\\ &= E(\xi'E(\xi|\mathscr{C}')|\mathscr{C})(\cdot).\end{aligned}$$

而由定理 (1) 及其后的注知此式右端 a.e. 等于 (2) 式右端, 故定理 2 获证.

二、条件分布. 设 $\xi_T=\{\xi_t,t\in T\}$ 是 $(\Omega,\mathscr{A},P)$ 上的随机函数, 为要研究 ξ_T 中某些随机变量的性质 (例如它们的函数的条件数学期望), 显然并不需要考虑整个 σ-代数 $\mathscr{A}$, 只需要考虑 ξ_T 所生成的子 σ-代数 $\sigma(\xi_T)$ 即可 (参看 §5.3). 即要研究随机函数 ξ_T, 只需要研究概率场 $(\Omega,\sigma(\xi_T),P)$; 要研究与 ξ_T 有关的集的条件概率, 只需要研究 $P(A|\mathscr{C}),A\in\sigma(\xi_T)$. 因此引出下列条件分布的概念.

定义 2　设 $(\Omega,\mathscr{A},P)$ 是概率空间,$\xi_T=\{\xi_t,t\in T\}$ 是其上的随机函数,$\mathscr{C}$ 是 $\mathscr{A}$ 的子 σ-代数. 如果函数 $P^{\mathscr{C}}(\omega,A),(\omega,A)\in\Omega\times\sigma(\xi_T)$ 满足以下两条件, 则称 $P^{\mathscr{C}}(\omega,A),(\omega,A)\in\Omega\times\sigma(\xi_T)$ 为ξ_T **在 $\mathscr{C}$ 之下 (关于 P) 的条件分布**.($\mathscr{C}$ 与 P 明确时, 简称为 ξ_T 的**条件分布**.)

i) 对一切 $A\in\sigma(\xi_T), P^{\mathscr{C}}(\cdot,A)$ 是 $\mathscr{C}$-可测函数, 对一切 $\omega\in\Omega, P^{\mathscr{C}}(\omega,\cdot)$ 是 $\sigma(\xi_T)$ 上的概率;

ii) 对一切 $A\in\sigma(\xi_T), P^{\mathscr{C}}(\cdot,A)$ 是 A 在 $\mathscr{C}$ 之下关于 P 的条件概率.

注意, 定义 2 中的 $\mathscr{C}$ 未必是 $\sigma(\xi_T)$ 的子 σ-代数. 如果 $\mathscr{C}\subset\sigma(\xi_T)$, 则条件分布就是 $(\Omega,\sigma(\xi_T),P)$ 上在 $\mathscr{C}$ 之下的正则条件概率.

容易看到, 如果 ξ,ξ' 是 $\sigma(\xi_T)$-可测的, 则将定理 1,2 中正则条件概率 $P^{\mathscr{C}}(\omega,A)$ 换作 ξ_T 的条件分布时定理 1,2 仍成立.

除了条件分布的概念而外, 我们还可以比照分布律的概念引进混合条件分布的概念.

三、混合条件分布

定义 3 设 $(\Omega,\mathscr{A},P)$ 是概率空间, $\xi_T=\{\xi_t,t\in T\}$ 是 T 上的随机函数,$\mathscr{C}$ 是 $\mathscr{A}$ 的子 σ-代数, $P^{\mathscr{C}}_{\xi_T}(\omega,B),(\omega,B)\in\Omega\times\mathscr{B}^T$(当 ξ_T 是复随机函数时以 $\mathscr{B}^T_z$ 代 $\mathscr{B}^T$) 满足

i) 对一切 $B\in\mathscr{B}^T, P^{\mathscr{C}}_{\xi_T}(\cdot,B)$ 是 $\mathscr{C}$-可测函数, 对一切 $\omega\in\Omega, P^{\mathscr{C}}_{\xi_T}(\omega,\cdot)$ 是 $\mathscr{B}^T$ 上的概率;

ii) 对任一 $B\in\mathscr{B}^T$, 存在一 $N\in\mathscr{C},P(N)=0$, 使

$$P^{\mathscr{C}}_{\xi_T}(\omega,B)=P(\{\xi_T\in B\}|\mathscr{C})(\omega),\omega\in N^c,$$

其中 $P(\{\xi_T\in B\}|\mathscr{C})$ 是 $\{\xi_T\in B\}$ 在 $\mathscr{C}$ 之下的条件概率, 则称 $P^{\mathscr{C}}_{\xi_T},(\omega,B),(\omega,B)\in\Omega\times\mathscr{B}^T$ 为ξ_T **在 $\mathscr{C}$ 之下 (关于 P) 的混合条件分布.**($\mathscr{C}$ 与 P 明确时, 简称为ξ_T **的混合条件分布.**)

如果 T 是有限集, 即 $\xi_T=\{\xi_1,\cdots,\xi_n\},\xi_T$ 是实的, $P^{\mathscr{C}}_{\xi_T}(\omega,B)$ 是 ξ_T 在 $\mathscr{C}$ 之下关于 P 的混合条件分布, 令

$$F^{\mathscr{C}}_{\xi_T}(\omega,x)=P^{\mathscr{C}}_{\xi_T}(\omega,(-\infty,x)),x=(x_1,\cdots,x_n)\in R^{(n)}.$$

则称此函数为 ξ_T 在 $\mathscr{C}$ 之下 (关于 P) 的条件分布函数.

为了避免繁琐的叙述在本节的余下部分均假设 ξ_T 为实随机函数.

ξ_T 的条件分布及混合条件分布的重要性在于: 当被积函数仅依赖于 ξ_T 时, 它具有与正则条件概率相同的积分性质. 即

定理 3 设 $g(x_T)$ 是 ξ_T 的值空间 R^T 上的 Borel 函数且 $Eg(\xi_T)$ 存在, 若 ξ_T 在 $\mathscr{C}$ 之下的条件分布 (混合条件分布) 存在, 则

$$E(g(\xi_T)|\mathscr{C})(\omega)=\int_\Omega g(\xi_T(\omega'))P^{\mathscr{C}}(\omega,d\omega')$$
$$\left(=\int_{R^T}g(x_T)P^{\mathscr{C}}_{\xi_T}(\omega,d_{x_T})\right),\omega\in N^c,\tag{3}$$

其中 $N \in \mathscr{C}, P(N)=0$.

证　由条件分布 (混合条件分布) 的定义知 (3) 对 $g=\chi_B, B \in \mathscr{B}^T$ 是成立的. 然后和通常一样的方法可证 (3) 对 g 是 Borel 简单函数、非负可测函数及满足定理条件的一般 Borel 函数都成立. □

以下几小节. 我们将要讨论条件分布及混合条件分布的存在问题. 我们将要证明当 T 是可数集时 ξ_T 的混合条件分布总是存在的. 在一般情况下, 这也就够用了, 因为关于 $\sigma(\xi_T)$ 可测的任一函数 f, 总存在 ξ_T 的可数子族 $\xi_{T_c}=\{\xi_{t_k,k\geqslant 1 t_k\in T_c}\}$, 使得 f 关于 $\sigma(\xi_{T_c})$ 可测. 当涉及到讨论 f 的条件期望时, 只要 ξ_{T_c} 在 $\mathscr{C}$ 之下的条件分布或混合条件分布存在就可以了.

为了搞清楚 ξ_T 在 $\mathscr{C}$ 之下条件分布与混合条件分布的存在性之间的关系, 我们来证明

定理 4　若 ξ_T 在 $\mathscr{C}$ 之下的条件分布存在, 则 ξ_T 有一在 $\mathscr{C}$ 之下的混合条件分布. 反之, 若 ξ_T 的值域是一 Borel 集, 且 ξ_T 在 $\mathscr{C}$ 之下的混合条件分布存在, 则 ξ_T 有一在 $\mathscr{C}$ 之下的条件分布.

注　所谓 ξ_T 的值域即 $B(\xi_T) \hat{=} \{x_T : \{\xi_T(\omega)=x_T\} \neq \varnothing\}$.

证　若 ξ_T 在 $\mathscr{C}$ 之下的条件分布 $P^{\mathscr{C}}(\omega, A), (\omega, A) \in \Omega \times \sigma(\xi_T)$ 存在, 则令

$$P^{\mathscr{C}}_{\xi_T}(\omega, B)=P^{\mathscr{C}}(\omega, \{\xi_T \in B\}), (\omega, B) \in \Omega \times \mathscr{B}^T.$$

显然 $P^{\mathscr{C}}_{\xi_T}(\omega, B)$ 是 ξ_T 在 $\mathscr{C}$ 之下的混合条件分布. 反之, 因为可能存在两个不同的 $B_k \in \mathscr{B}^T, k=1,2$, 使 $A=\{\xi_T \in B_k\}, k=1,2$, 于是 $P^{\mathscr{C}}(\omega, A)$ 就不能由 $P^{\mathscr{C}}_{\xi_T}(\omega, B_k)$ 唯一确定. 但若 $B(\xi_T)$ 是 Borel 集 (即 $B(\xi_T) \in \mathscr{B}^T$) 且 $P^{\mathscr{C}}_{\xi_T}(\omega, B)$ 是 ξ_T 在 $\mathscr{C}$ 之下的混合条件分布, 则由混合条件分布的定义知

$$\begin{aligned} P^{\mathscr{C}}_{\xi_T}(\omega, [B(\xi_T)]^c) &= P(\xi_T \in [B(\xi_T)]^c | \mathscr{C}) \\ &= P(\varnothing | \mathscr{C}) = P^{\mathscr{C}}_{\xi_T}(\omega, \varnothing)=0, \omega \in N^c, \end{aligned}$$

其中 $N \in \mathscr{C}, P(N)=0$.

对于任意的 $A \in \sigma(\xi_T)$ 来说, 若有 $B_1, B_2 \in \mathscr{B}^T$ 使 $A=\{\xi_T \in B_k\}, k=1,2$, 则 $B_1 \cap B_2^c \subset [B(\xi_T)]^c, B_2 \cap B_1^c \subset [B(\xi_T)]^c$, 因而由 (3) 及 $P^{\mathscr{C}}_{\xi_T}(\omega, B)$ 是 $\mathscr{B}^T$ 上的概率知

$$0 \leqslant P^{\mathscr{C}}_{\xi_T}(\omega, B_1 \cap B_2^c) \leqslant P^{\mathscr{C}}_{\xi_T}(\omega, [B(\xi_T)]^c)=0, \omega \in N^c.$$

同理有 $P^{\mathscr{C}}_{\xi_T}(\omega, B_2 \cap B_1^c)=0, \omega \in N^c$. 因此

$$P^{\mathscr{C}}_{\xi_T}(\omega, B_1)=P^{\mathscr{C}}_{\xi_T}(\omega, B_2)=P^{\mathscr{C}}_{\xi}(\omega, B_1 \cap B_2), \omega \in N^c.$$

对每一 $A \in \sigma(\xi_T)$, 选定 $B \in \mathscr{B}^T$ 使 $A = \{\xi_T \in B\}$, 而用等式

$$P^{\mathscr{C}}(\omega, A) = \begin{cases} P^{\mathscr{C}}_{\xi_T}(\omega, B), & \omega \in N^c, \\ P^{\mathscr{C}}_{\xi_T}(\omega_0, B), & \omega \in N \end{cases} \tag{4}$$

(其中 ω_0 是 N^c 中任意选定的元素) 定义的 $P^{\mathscr{C}}(\omega, A)$ 对每一 $\omega \in \Omega$ 及 $A \in \sigma(\xi_T)$ 是唯一确定的, 且对每一 $A \in \sigma(\xi_T)$ 是 $\mathscr{C}$ 可测函数, 对每一 $\omega \in \Omega, P^{\mathscr{C}}(\omega, \cdot)$ 是 $\sigma(\xi_T)$ 上的概率. 事实上, 非负正规显然, 只需证 σ-可加性. 若 $A_n = \{\xi_T \in B_n\}, n = 1, 2, \cdots$ 两两不交, 则 $A_n = \{\xi_T \in B_n \cap B(\xi_T)\}$, 令 $B'_n = B_n \cap B(\xi_T)$, 则 $B'_n, n = 1, 2, \cdots$ 两两不交. 因而

$$\begin{aligned} P^{\mathscr{C}}(\omega, \textstyle\sum A_n) &= P^{\mathscr{C}}(\omega, \{\xi_T \in \textstyle\sum B'_n\}) \\ &= P^{\mathscr{C}}_{\xi_T}(\omega, \textstyle\sum B'_n) = \textstyle\sum P^{\mathscr{C}}_{\xi_T}(\omega, B'_n) \\ &= \textstyle\sum P^{\mathscr{C}}(\omega, A_n), \end{aligned}$$

(当 $\omega \in N$ 时 $P^{\mathscr{C}}_{\xi_T}(\omega, \sum B'_n)$ 及 $P^{\mathscr{C}}_{\xi_T}(\omega, B'_n)$ 中的 ω 以 ω_0 代替) 故 $P^{\mathscr{C}}(\omega, \cdot)$ 是 $\sigma(\xi_T)$ 上的概率.

再由 $P^{\mathscr{C}}_{\xi_T}(\cdot, B)$ 与 $A = \{\xi_T \in B\}$ 在 $\mathscr{C}$ 之下的条件概率几乎相等, 因而由 (4) 中第一式知 $P^{\mathscr{C}}(\omega, A)$ 对每一 $A \in \sigma(\xi_T)$ 来说是 A 在 $\mathscr{C}$ 之下的条件概率. 故 $P^{\mathscr{C}}(\omega, A)$ 是 ξ_T 在 $\mathscr{C}$ 之下的条件分布. □

当 ξ_T 是复随机函数时, 只需以 $\mathscr{B}^T_z$ 代替 $\mathscr{B}^T$, 定理仍然成立.

5.5.2 混合条件分布的存在及正则化定理 (有限维情形)

定理 5 设 $\xi = (\xi_1, \cdots, \xi_n)$ 是 $(\Omega, \mathscr{A}, P)$ 上的 n 维实随机变量, $\mathscr{C}$ 为 $\mathscr{A}$ 的任一子 σ-代数, 则 ξ 在 $\mathscr{C}$ 之下 (关于 P) 的混合条件分布是存在的, 若 ξ 的值域是 $R^{(n)}$ 中的 Borel 集, 则 ξ 在 $\mathscr{C}$ 之下的条件分布存在.

证 定理的后一部分由前一部分及定理 4 立刻得出, 故只需证定理的前一部分.

证明的构思基于下列事实: $R^{(n)}$ 中有理坐标点是可数的, 并且在其中稠密.

以 x, x' 表 $R^{(n)}$ 中的点, 以 r, r' 表其中各个坐标皆为有理数的点 (称为有理点), 以 M, N(或带有足码) 表 $\mathscr{C}$ 中的零概率集.

选定一个在 $\mathscr{C}$ 之下的条件概率 $P(A|\mathscr{C})(\omega), (\omega, A) \in \Omega \times \mathscr{A}$. 考虑函数族 $P(\{\xi < r\}|\mathscr{C})(\cdot)$, 根据条件期望及条件概率的性质 (5.2.2 节定理 1,2 及定理 1′) 不难得出对于任意有理点 r 及 r', 以及 $\tilde{r}_k = (r_1, \cdots, r_{k-1}, r_{k+1}, \cdots, r_n)$ (其中 $k = 1, \cdots, n, r_j$ 为任意有理数,$j = 1, \cdots, n$) 存在 $\mathscr{C}$ 中零概率集 $N_{rr'}, M_{rr'}, N_r, M_{\tilde{r}k}, M$, 使得

$$\Delta_{r'r} P(\{\xi < x\}|\mathscr{C})(\omega) \geqslant 0, \omega \in N^c_{rr'}, r' \geqslant r;$$

$$0 \leqslant P(\{\xi < r\}|\mathscr{C})(\omega) \leqslant P(\{\xi < r'\}|\mathscr{C})(\omega) \leqslant 1, \omega \in M_{rr'}^c, r' \geqslant r;$$

$$P\left(\left\{\xi < r - \frac{1}{m}\right\}\middle|\mathscr{C}\right)(\omega) \to P(\{\xi < r\}|\mathscr{C})(\omega), m \to \infty, \omega \in N_r^c;$$

$$\lim_{m\to\infty} P(\{\xi < m\}|\mathscr{C})(\omega) = 1, \omega \in M^c;$$

$$\lim_{m\to\infty} P(\{\xi < \mathrm{S_{km}}\}|\mathscr{C})(\omega) = 0, \omega \in M_{\tilde{r}_k}^c,$$

其中 $\mathrm{S_{km}} = (r_1, \cdots r_{k-1}, -m, r_{k+1}, \cdots, r_n)$. 令

$$N = \bigcup_{r,r'} \left(N_{rr'} \cup M_{rr'} \cup N_r \cup \left(\bigcup_{k=1}^n M_{\tilde{r}_k} \right) \right) \cup M.$$

则 N 为 $\mathscr{C}$ 中零概率集, 易证当 $\omega \in N^c$ 时, 有

$$\begin{cases} \Delta_{r'r} P(\{\xi < x\}|\mathscr{C})(\omega) \geqslant 0; \\ \lim\limits_{r \nearrow r'} P(\{\xi < r\}|\mathscr{C})(\omega) = P(\{\xi < r'\}|\mathscr{C})(\omega); \\ \lim\limits_{r\to\infty} P(\{\xi < r\}|\mathscr{C})(\omega) = 1; \\ \lim\limits_{r_k\to-\infty} P(\{\xi < r\}|\mathscr{C})(\omega) = 0, \end{cases} \tag{5}$$

其中 $r = (r_1, \cdots, r_n)$, 及

$$P(\{\xi < r\}|\mathscr{C})(\omega) \leqslant P(\{\xi < r'\}|\mathscr{C})(\omega), r \leqslant r'. \tag{6}$$

定义

$$F^{\mathscr{C}}(\omega, r) = \begin{cases} P(\{\xi < r\}|\mathscr{C})(\omega), \omega \in N^c, \\ P(\{\xi < r\}|\mathscr{C})(\omega_0), \omega \in N, \end{cases} \tag{7}$$

其中 ω_0 是 N^c 中的任一取定的元.

其次, 对每一 $x \in R^{(n)}$ 定义

$$F^{\mathscr{C}}(\omega, x) = \lim_{r\uparrow x} F^{\mathscr{C}}(\omega, r). \tag{8}$$

由于由 (7) 定义的 $F^{\mathscr{C}}(\omega, r)$ 在有理点上非降左连续, 因而 (7) 与 (8) 定义的 $F^{\mathscr{C}}(\omega, x)$ 在 x 为有理点时一致.

由 (5) 知 $F^{\mathscr{C}}(\omega, x)$ 在有理点上满足概率分布函数的条件. 由 (5),(6) 及 (8) 推知对一切 $x \in R^{(n)}, F^{\mathscr{C}}(\omega, x)$ 满足概率分布函数的条件, 即对每一 $\omega \in \Omega, F^{\mathscr{C}}(\omega, x)$ 是 $R^{(n)}$ 上的概率分布函数. 因而由第 2 章 §2.3 定理 3 知, 对每一 $\omega \in \Omega$, 存在 $(R^{(n)}, \mathscr{B}^{(n)})$ 上的概率测度 $P^{\mathscr{C}}(\omega, \cdot)$ 使得

$$P^{\mathscr{C}}(\omega, (-\infty, x)) = F^{\mathscr{C}}(\omega, x), \text{对一切} x \in R^{(n)}.$$

现在我们来证明 $P^{\mathscr{C}}(\omega, B), (\omega, B) \in \Omega \times \mathscr{B}^{(n)}$ 是 ξ 在 $\mathscr{C}$ 之下的混合条件分布 (即 $P_{\xi}^{\mathscr{C}}(\omega, B)$). 为此只需证明对每一 $B \in \mathscr{B}^{(n)}, P^{\mathscr{C}}(\cdot, B)$ 是 $\mathscr{C}$ 可测函数及对每一 $B \in \mathscr{B}^{(n)}$

$$P^{\mathscr{C}}(\cdot, B) = P(\{\xi \in B\}|\mathscr{C}), \text{a.e.}$$

为证此事实, 令

$$\pi \hat{=} \{B = (-\infty, r) : r\text{为}R^{(n)}\text{中的有理点}\},$$

$$\lambda \hat{=} \{B : B \in \mathscr{B}^{(n)}, P^{\mathscr{C}}(\cdot, B) \in \mathscr{C}\text{且}P^{\mathscr{C}}(\cdot, B) = P(\{\xi \in B\}|\mathscr{C}), \text{a.e.}\}.$$

首先,$\lambda \supset \pi, \pi$ 是 π-系且 π 上最小 λ-系 $\lambda(\pi) = \sigma(\pi) = \mathscr{B}^{(n)}$.

再者, 应用概率的连续性及条件概率的性质易证 λ 为 λ-系. 于是定理获证. □

注 当 ξ 是 n 维复随机变量时定理仍然成立 (将定理叙述中的 $R^{(n)}$ 代之以 $Z^{(n)}$). 事实上, 令 $\xi_j = \xi_{j1} + i\xi_{j2}, \xi_{j1}, \xi_{j2}$ 为实随机变量, $j = 1, \cdots, n$. 此时考虑 $\tilde{\xi} = \{\xi_{11}, \xi_{12}, \cdots, \xi_{n1}, \xi_{n2}\}, \tilde{\xi}$ 的混合条件分布 $P_{\tilde{\xi}}^{\mathscr{C}}$ 存在. 由于 $\mathscr{B}_z^{(n)}$ 与 $\mathscr{B}^{(2n)}$ 存在一一对应, 并且保持集合运算关系, 故对任一 $B_z^{(n)} \in \mathscr{B}_z^{(n)}$, 令

$$P_{\xi}^{\mathscr{C}}(\omega, B_z^{(n)}) \hat{=} P_{\tilde{\xi}}^{\mathscr{C}}(\omega, B^{(2n)}),$$

其中

$$B^{(2n)} = \{(x_1, y_1, \cdots, x_n, y_n) : (x_1 + iy_1, \cdots, x_n + iy_n) \in B_z^{(n)}\} \in \mathscr{B}^{(2n)},$$

并且还有

$$\{\xi \in B_z^{(n)}\} = \{\tilde{\xi} \in B^{(2n)}\}.$$

因此 $P_{\xi}^{\mathscr{C}}$ 是 ξ 在 $\mathscr{C}$ 之下的混合条件分布.

由定理 5 不难得出在一定意义下正则条件概率是存在的. 即

定理 6 对于概率空间 $(R^{(n)}, \mathscr{B}^{(n)}, P)$ 及 $\mathscr{B}^{(n)}$ 的任一子 σ-代数 $\mathscr{C}$, 永远存在在 $\mathscr{C}$ 之下的正则条件概率 $P^{\mathscr{C}}(x, A), (x, A) \in R^{(n)} \times \mathscr{B}^{(n)}$.

证 令

$$\xi_k(x) = x_k, k = 1, \cdots, n, x = (x_1, \cdots, x_n) \in R^{(n)},$$

显然 $\xi = (\xi_1, \cdots, \xi_n)$ 是 $(R^{(n)}, \mathscr{B}^{(n)}, P)$ 上的一个 n 维随机变量, 而 ξ 的值域为 $R^{(n)} \in \mathscr{B}^{(n)}$, 故由定理 5 知 ξ 在 $\mathscr{C}$ 之下的条件分布 $P^{\mathscr{C}}(x, A), (x, A) \in R^{(n)} \times \sigma(\xi)$ 存在, 但易见 $\sigma(\xi) = \mathscr{B}^{(n)}$, 故由条件分布的定义知 $P^{\mathscr{C}}(x, A), (x, A) \in R^{(n)} \times \mathscr{B}^{(n)}$ 即为 $\mathscr{C}$ 之下的正则条件概率. □

在 §5.3 例 4 中, 如果取 $(R^{(n+m)}, \mathscr{B}^{(n+m)})$ 为基本空间, $\zeta = (\xi_1, \cdots, \xi_n, \eta_1, \cdots, \eta_m)$ 为样本函数, 即对任一 $\omega = (x_1, \cdots, x_n, y_1, \cdots, y_m) \in R^{(n+m)}$,

$$\xi_k(\omega) = x_k, k = 1, \cdots, n,$$

$$\eta_j(\omega) = y_j, j = 1, \cdots, m.$$

若令 $\xi = (\xi_1, \cdots, \xi_n)$, 且对 $\omega = (x_1, \cdots, x_n, y_1, \cdots, y_m) \in R^{(n+m)}, B^{(n)} \in \mathscr{B}^{(n)}$ 令

$$P^{\sigma(\eta)}(\omega, \{\xi \in B^{(n)}\}) = \begin{cases} P(\xi \in B^{(n)} | \eta_1 = y_1, \cdots, \eta_m = y_m), \text{若} y \notin N \cup \{p_\eta(y) = 0\}, \\ P(\xi \in B^{(n)} | \eta_1 = y_1^{(0)}, \cdots, \eta_m = y_m^{(0)}), \text{若} y \in N \cup \{p_\eta(y) = 0\}, \end{cases}$$

其中 $y = (y_1, \cdots, y_m), y^{(0)} = (y_1^{(0)}, \cdots, y_m^{(0)}) \notin N \cup \{p_\eta(y) = 0\}$ 任意取定. 而 $P(\xi \in B^{(n)} | \eta_1 = y_1, \cdots, \eta_m = y_m)$ 由 §5.3(10) 式给定. 则 $P^{\sigma(\eta)}(\omega, \{\xi \in B^{(n)}\})$ 是 ξ 在 $\sigma(\eta)$ 下的条件分布.

定理 7　设 P 是 $(R^{(n)}, \mathscr{B}^{(n)})$ 上的概率, 则存在着 $\mathscr{B}^{(1)}$ 上的概率 P_1 及 $R^{(k-1)} \times \mathscr{B}_k$ 上的转移概率 $P_k(x_1, \cdots, x_{k-1}, B_k), (x_1, \cdots, x_{k-1}) \in R^{(k-1)}, B_k \in \mathscr{B}_k, k = 2, \cdots, n$, 使得对一切 $B^{(n)} \in \mathscr{B}^{(n)}$,

$$\begin{aligned} P(B^{(n)}) = & \int_{R_1} \int_{R_2} \cdots \int_{R_n} \chi_{B^{(n)}}(x_1, \cdots, x_{n-1}, x_n) \\ & \times P_n(x_1, \cdots, x_{n-1}, dx_n) \cdots P_2(x_1, dx_2) P_1(dx_1), \end{aligned} \tag{9}$$

即 P 由 $P_1(B_1), P_2(x_1, B_2), \cdots, P_n(x_1, \cdots, x_{n-1}, B_n), x_k \in R_k, B_k \in \mathscr{B}_k, k = 1, \cdots, n$ 通过 (9) 来决定.(这里 $R_k = R^{(1)}, \mathscr{B}_k = \mathscr{B}^{(1)}, k = 1, \cdots, n$.)

证　令

$$\mathscr{C}_{n-1} = \{B^{(n-1)} \times R_n : B^{(n-1)} \in \mathscr{B}^{(n-1)}\},$$

则由定理 6 知有一在 $\mathscr{C}_{n-1}$ 之下的正则条件概率 $P^{\mathscr{C}_{n-1}}(x, A), (x, A) \in R^{(n)} \times \mathscr{B}^{(n)}$ 存在, 由于对任意给定的 $(x_1, \cdots, x_{n-1}) \in R^{(n-1)}$, 集合 $\{(x_1, \cdots, x_{n-1})\} \times R_n$ 是 $\mathscr{C}_{n-1}$ 中的原子, 故由 §2.3 定理 5 知等式

$$P_n(x_1, \cdots, x_{n-1}, B_n) = P^{\mathscr{C}_{n-1}}\left(x_1, \cdots, x_n, \prod_{k=1}^{n-1} R_k \times B_n\right)$$

定义了 $R^{(n-1)} \times \mathscr{B}_n$ 上的函数, 而且由正则条件概率的定义易知它是 $(R^{(n-1)}, \mathscr{B}^{(n-1)})$ 到 $(R_n, \mathscr{B}_n)$ 上的转移概率.

我们把 P 在 $\mathscr{C}_{n-1}$ 上的限制看作 $(R^{(n-1)}, \mathscr{B}^{(n-1)})$ 上的概率, 记作 $P_{1,\cdots,n-1}$. 由于对任何 $B_k \in \mathscr{B}_k, k=1,\cdots,n$,

$$\begin{aligned}\int_{\prod_{k=1}^{n-1} B_k} P_n(x_1,\cdots,x_{n-1},B_n)dP_{1,\cdots,n-1} &= \int_{\prod_{k=1}^{n-1} B_k\times R_n} P^{\mathscr{C}_{n-1}}\left(x, \prod_{k=1}^{n-1} R_k \times B_n\right)dP \\ &= P\left(\left(\prod_{k=1}^{n-1} R_k \times B_n\right)\cap\left(\prod_{k=1}^{n-1} B_k \times R_n\right)\right) \\ &= P(B_1\times\cdots\times B_n).\end{aligned} \tag{10}$$

由测度扩张定理知 P 由 $P_n(x_1,\cdots,x_{n-1},B_n),(x_1,\cdots,x_{n-1},B_n)\in R^{(n-1)}\times\mathscr{B}_n$ 及 $P_{1,\cdots,n-1}$ 唯一决定. 对 $P_{1,\cdots,n-1}$ 重复上面的讨论并继续做下去即知 $\mathscr{B}^{(n)}$ 上的概率 P 由转移概率序列

$$P_1(B_1), P_2(x_1,B_2),\cdots,P_n(x_1,\cdots,x_{n-1},B_n), B_k\in\mathscr{B}_k, x_k\in R_k, k=1,\cdots,n$$

来决定. 其中

$$P_k(x_1,\cdots,x_{k-1},B_k) = P_{1,\cdots,k}^{\mathscr{C}_{k-1}}\left(x_1,\cdots,x_k,\prod_{j=1}^{k-1} R_j\times B_k\right), k=2,\cdots,n$$

是在 $\mathscr{C}_{k-1}=\{B^{(k-1)}\times R_k : B^{(k-1)}\in\mathscr{B}^{(k-1)}\}$ 之下关于 $P_{1,\cdots,k}$ 的正则条件概率, 而

$$P_{1,\cdots,n}(B^{(n)}) = P(B^{(n)}), B^{(n)}\in\mathscr{B}^{(n)},$$

$$P_{1,\cdots,k}(B^{(k)}) = P_{1,\cdots,k,k+1}(B^{(k)}\times R_{k+1}), B^{(k)}\in\mathscr{B}^{(k)}, k=1,\cdots,n-1.$$

以下往证 (9) 式成立. 使用归纳法. 由 (10) 及测度扩张定理知当 $n=2$ 时 (9) 式成立. 若 (9) 式对 $n=k-1$ 成立, 令证对 k 成立. 设 $B^{(k)}=B^{(k-1)}\times B_k$, 同证明 (10) 一样的方法可知对任何 $B^{(k-1)}\in\mathscr{B}^{(k-1)}, B_k\in\mathscr{B}_k$

$$\begin{aligned}P_{1,\cdots,k}(B^{(k-1)}\times B_k) &= \int_{B^{(k-1)}} P_k(x_1,\cdots,x_{k-1},B_k)dP_{1,\cdots,k-1} \\ &= \int_{R^{(k-1)}} \chi_{B^{(k-1)}}(x_1,\cdots,x_{k-1})P_k(x_1,\cdots,x_{k-1},B_k)dP_{1,\cdots,k-1},\end{aligned}$$

其中 $\chi_{B^{(k-1)}}(x_1,\cdots,x_{k-1})P_k(x_1,\cdots,x_{k-1},B_k),(x_1,\cdots,x_{k-1})\in R^{(k-1)}$ 是 $\mathscr{B}^{(k-1)}$ 可测函数. 故由 (9) 对 $(k-1)$ 成立及 §5.4 定理 3 知

$$P_{1,\cdots,k}(B^{(k-1)}\times B_k) = \int_{R_1}\int_{R_2}\cdots\int_{R_{k-1}} \chi_{B^{(k-1)}}(x_1,\cdots,x_{k-1})$$

$$\times P_k(x_1,\cdots,x_{k-1},B_k)P_{k-1}(x_1,\cdots,x_{k-2},dx_{k-1})\cdots$$

$$P_2(x_1,dx_2)P_1(dx_1)$$

$$=\int_{R_1}\int_{R_2}\cdots\int_{R_{k-1}}\int_{R_k}\chi_{B^{(k-1)}\times B_k}(x_1,\cdots,x_k)$$

$$\times P_k(x_1,\cdots,x_{k-1},dx_k)P_{k-1}(x_1,\cdots,x_{k-2},dx_{k-1})\cdots$$

$$P_2(x_1,dx_2)P_1(dx_1).$$

因而 (9) 式对 $n=k$ 且 $B^{(k)}=B^{(k-1)}\times B_k, B^{(k-1)}\in\mathscr{B}^{(k-1)}, B_k\in\mathscr{B}_k$ 时成立. 由测度扩张定理及 (9) 式右端是一个测度 (§5.4 定理 2) 知 (9) 对 $n=k$ 时成立. 因此 (9) 对一切自然数 n 成立. 定理获证. □

5.5.3 Колмогоров 和谐定理

我们先引入有限维概率和谐的概念.

定义 4　设 T 为一无穷集 (可数或不可数), 对每一 $t\in T$, 有一可测空间 $(\Omega_t,\mathscr{A}_t)$ 对应, 若对 T 的某一有限子集类中的每一 T_N, 有一概率场 $(\Omega^{T_N},\mathscr{A}^{T_N}, P^{T_N})$, 其中 $\Omega^{T_N}=\prod\limits_{t\in T_N}\Omega_t, \mathscr{A}^{T_N}=\prod\limits_{t\in T_N}\mathscr{A}_t$, 且它们具有下列性质: 对于 T 的这个有限子集类中的任何两个元 $T_N, T'_N, T_N\subset T'_N$ 及任一 $A^{T_N}\in\mathscr{A}^{T_N}$,

$$P^{T_N}(A^{T_N})=P^{T'_N}\left(A^{T_N}\times\prod_{t\in T'_N-T_N}\Omega_t\right),$$

则称该概率场是和谐的.

定理 8(Колмогоров和谐定理)　设 T 为一无穷集, 对 T 的任一有限子集 T_N 来说, 有一概率场 $(R^{T_N},\mathscr{B}^{T_N},P^{T_N})$, 其中 $R^{T_N}=\prod\limits_{t\in T_N}R_t, \mathscr{B}^{T_N}=\prod\limits_{t\in T_N}\mathscr{B}_t$, 每一 $(R_t,\mathscr{B}_t)$ 是一维 Borel 空间. 如果这些概率场是和谐的, 则在 $(R^T,\mathscr{B}^T)$ 上有唯一的概率 P^T 存在, 使得对任何 $T_N\subset T$ 及以 $B^{T_N}\in\mathscr{B}^{T_N}$ 为底的 Borel 柱体 $B^{T_N}\times\prod\limits_{t\in T-T_N}R_t$, 有

$$P_T\left(B^{T_N}\times\prod_{t\in T-T_N}R_t\right)=P^{T_N}(B^{T_N}).\tag{11}$$

证　i) 当 T 是可数无穷集时, 不妨设 $T=\{1,2,\cdots\}$. 于是概率场 $(R^{(n)},\mathscr{B}^{(n)}, P^{(n)}), n=1,2,\cdots$ 是和谐的. 因此由定理 7 的证明知它们决定一个转移概率序列

$$P^{(1)}(B_1),P_2(x_1,B_2),\cdots,P_n(x_1,\cdots,x_{n-1},B_n),\cdots,x_k\in R_k,B_k\in\mathscr{B}_k,$$

其中

$$P_n(x_1,\cdots,x_{n-1},B_n)=(P^{(n)})^{\mathscr{C}_n-1}(x_1,\cdots,x_n,R^{(n-1)}\times B_n),n=2,\cdots.$$

而 $(P^{(n)})^{\mathscr{C}_n-1}$ 是在 $\mathscr{C}_{n-1}$ 下关于 $P^{(n)}$ 的正则条件概率. 于是由 §5.4 定理 5 知在 $(R^{(\infty)},\mathscr{B}^{(\infty)})$ 上有唯一确定的概率 $P^{(\infty)}$ 存在, 使得对任一 $B^{(n)}\in\mathscr{B}^{(n)}$,

$$P^{(\infty)}\left(B^{(n)}\times\prod_{k=n+1}^{\infty}R_k\right)=\int_{R_1}\int_{R_2}\cdots\int_{R_n}\chi_{B^{(n)}}(x_1,\cdots,x_n)\times P_n(x_1,\cdots,x_{n-1},dx_n)\cdots P_2(x_1,dx_2)P^{(1)}(dx_1). \quad (12)$$

再由定理 7 知 (12) 式的右端即 $P^{(n)}(B^{(n)})$.

对这种情形, 还需证明对任意 $T_N\subset T,T_N$ 有限集有

$$P^{(\infty)}\left(B^{T_N}\times\prod_{t\in T-T_N}R_n\right)=P^{T_N}(B^{T_N}),B^{T_N}\in\mathscr{B}^{T_N}.$$

事实上, 存在 m 使 $T_N\subset\{1,\cdots,m\}$, 于是

$$B^{T_N}\times\prod_{n\in T-T_N}R_n=\left(B^{T_N}\times\prod_{n\in\{1,\cdots,m\}-T_N}R_n\right)\times\prod_{n=m+1}^{\infty}R_n.$$

因而

$$P^{(\infty)}\left(B^{T_N}\times\prod_{n\in T-T_N}R_n\right)=P^{(m)}\left(B^{T_N}\times\prod_{n\in\{1,\cdots,m\}-T_N}R_n\right)=P^{(T_N)}(B^{T_N}).$$

最后一步是由于 $P^{T_N},T_N\subset T$ 的和谐性.

ii) 设 T 是不可数集, 任取 T 的一个可数无穷子集 T_c, 则由 i) 知在 $\mathscr{B}^{T_c}$ 上有唯一的概率 P^{T_c} 存在, 使得对任何以 $B^{T_N}(T_N\subset T_c)$ 为底的 Borel 柱体 $B^{T_N}\times\prod\limits_{t\in T_c-T_N}R_t$ 都有

$$P^{T_c}\left(B^{T_N}\times\prod_{t\in T_c-T_N}R_t\right)=P^{T_N}(B^{T_N}). \quad (13)$$

由第 4 章 §4.3 引理 1 知对任一 $B^T\in\mathscr{B}^T$, 必有一 T 的可数子集 T_c 使得 $B^T=B^{T_c}\times\prod\limits_{t\in T-T_c}R_t$. 令

$$P^T(B^T)\hat{=}P^{T_c}(B^{T_c}). \quad (14)$$

几乎逐字逐句地重复第四章 §4.3 定理 2 的证明, 可知 (14) 式所定义的 P^T 是 $(R^T,\mathscr{B}^T)$ 上唯一满足 (11) 的概率. □

设 T 为任一指标集, $(R^T,\mathscr{B}^T)$ 为定理 8 中的可测空间, P^T 为 $\mathscr{B}^T$ 上的概率, 显然由等式

$$\tilde{\xi}_t(x^T)=x_t, x^T=\{x_t,t\in T\}\in R^T, t\in T \tag{15}$$

定义的 R^T 上的函数族 $\tilde{\xi}_T=\{\tilde{\xi}_t,t\in T\}$ 是概率空间 $(R^T,\mathscr{B}^T,P^T)$ 上的随机函数, 且 $P^T(\tilde{\xi}_T\in B^T)=P^T(B^T), B^T\in\mathscr{B}^T$. 对于 T 的任一有限子集 T_N 来说, 我们称

$$P^{T_N}(B^{T_N})=P^T\left(B^{T_N}\times\prod_{t\in T-T_N}R_t\right), B^{T_N}\in\mathscr{B}^{T_N} \tag{16}$$

是 P^T**在**$\mathscr{B}^{T_N}$**上的边缘概率 (或在 $\mathscr{B}^{T_N}$ 上的射影)**. 显然随机变量 $\tilde{\xi}_{T_N}=\{\tilde{\xi}_{t_1},\cdots,\tilde{\xi}_{t_N}\}$ (其中 $T_N=\{t_1,\cdots,t_N\}$) 的分布律 $P_{\tilde{\xi}_{T_N}}$ 是 P^T 在 $\mathscr{B}^{T_N}$ 上的边缘概率.

设 $\xi_T=\{\xi_t,t\in T\}$ 是任一给定的概率空间 $(\Omega,\mathscr{A},P)$ 上的随机函数. 定义

$$P_{\xi_T}(B^T)=P(\xi_T\in B^T), B^T\in\mathscr{B}^T.$$

于是 P_{ξ_T} 是 $(R^T,\mathscr{B}^T)$ 上的概率, 称为**随机函数 ξ_T 的分布**. 若令 $\{P_{\xi_{T_N}},T_N\subset T\}$ 是 ξ_T 的有限维分布族, 则它们是和谐的, 因而 P_{ξ_T} 由 $\{P_{\xi_{T_N}},T_N\subset T\}$ 唯一决定.

从概率角度 (而不是从可测函数的角度) 来讨论随机函数 $\xi_T=\{\xi_t,t\in T\}$ 的性质时, ξ_T 与它所决定的概率空间 $(R^T,\mathscr{B}^T,P^T)$(其中 $P^T=P_{\xi_T}$) 上的随机函数 $\tilde{\xi}_T$ 可以看成相同的 (即在同分布的意义下相同). 因此在只讨论有关 ξ_T 的概率分布的性质时,$(R^T,\mathscr{B}^T,P^T)$ 有其普遍的意义. 故我们给出下列的

定义 5　设 $\xi_T=\{\xi_t,t\in T\}$ 是 $(\Omega,\mathscr{A},P)$ 上的随机函数, 则称 $(R^T,\mathscr{B}^T,P_{\xi_T})$ 为 ξ_T**的概率样本空间**,而称 $(R^T,\mathscr{B}^T)$ 为 ξ_T**的样本空间**, 称由 (15) 定义的 $\tilde{\xi}_T=\{\tilde{\xi}_t,t\in T\}$ 为 ξ_T**的表现.**

定义 6　设 T 是任一指标集, 对任意 $T_N=\{t_1,\cdots,t_N\}, F_{t_1,\cdots,t_N}$ 是 $(R^{T_N},\mathscr{B}^{T_N})$ 上的概率分布函数且满足

i) 对 $1,2,\cdots,N$ 的任一排列 $\alpha_1,\cdots,\alpha_N$ 有

$$F_{t_1,\cdots,t_N}(x_1,\cdots,x_n)=F_{t_{\alpha_1},\cdots,t_{\alpha_N}}(x_{\alpha_1},\cdots,x_{\alpha_N}), (x_1,\cdots,x_N)\in R^{T_N}; \tag{17}$$

ii) 若 $m<n$, 则

$$F_{t_1,\cdots,t_m}(x_1,\cdots,x_m)=\lim_{x_{m+1},\cdots,x_n\to\infty}F_{t_1,\cdots,t_n}(x_1,\cdots,x_n), (x_1,\cdots,x_m)\in R^{T_m}$$

则称分布函数族 $\{F_{t_1,\cdots,t_N}:(t_1,\cdots,t_N)\subset T\}$ 是和谐的.

推论 1　设 T 是任一指标集,$\{F_{t_1,\cdots,t_N}:(t_1,\cdots,t_N)\subset T\}$ 是和谐的分布函数族, 则在 $(R^T,\mathscr{B}^T)$ 上有唯一的概率 P^T 存在, 使得由 (15) 定义的随机函数 $\tilde{\xi}_T$ 的任何有限维分量 $(\tilde{\xi}_{t_1},\cdots,\tilde{\xi}_{t_N})$ 的分布函数是 $F_{t_1,\cdots,t_N}$.

证 由 $\{F_{t_1,\cdots,t_N}: T_N=(t_1,\cdots,t_N)\subset T\}$ 的和谐性及第 2 章 2.3.2 定理 2 易知由它们决定的概率场 $(R^{T_N},\mathscr{B}^{T_N},P^{T_N}),T_N\subset T$ 是和谐的. 因而由定理 8 知有唯一的概率场 $(R^T,\mathscr{B}^T,P^T)$ 存在, 使 P^{T_N} 是 P^T 在 $\mathscr{B}^{T_N}$ 上的边缘概率. 再由定理 8 后面的注解知 $\tilde{\xi}_{T_N}$ 的概率分布是 P^{T_N}, 因而分布函数是 $F_{t_1,\cdots t_N}$, 其中 $(t_1,\cdots,t_N)=T_N$. □

定理 8 及推论 1 说明只要给了和谐的分布函数族 $\{F_{T_N},T_N\subset T\}$(或和谐的概率空间族 $(R^{T_N},\mathscr{B}^{T_N},P^{T_N}),T_N\subset T$), 就可以构造一个随机函数 ξ_T, 使得对 T 的任一有限子集 T_N,ξ_{T_N} 的分布函数 (或概率分布) 是 F_{T_N}(或 P^{T_N}). 这个结论是随机过程理论的基础之一.

例 1 称随机变量序列 $\xi_1,\xi_2,\cdots$ 为一马氏序列, 如果对任意 $B\in\mathscr{B}^{(1)}$, $(x_1,\cdots,x_{n-1})\in R^{(n-1)},n=2,3,\cdots$ 有

$$P(\xi_n\in B|\xi_1=x_1,\cdots,\xi_{n-1}=x_{n-1})=P(\xi_n\in B|\xi_{n-1}=x_{n-1}),\text{a.e.}$$

若给定一组转移概率序列 $P_1(B_1),B_1\in\mathscr{B}^{(1)},P_n(x_1,\cdots,x_{n-1},B_n),B_n\in\mathscr{B}^{(1)},n=2,3,\cdots$ 且对任何 $n,B_n\in\mathscr{B}^{(1)}$ 及任何 $(x_1,\cdots,x_{n-1})\in R^{(n-1)}$ 都有

$$P_n(x_1,\cdots,x_{n-1},B_n)=P_n(x_{n-1},B_n),$$

则有一定义在概率样本空间 $(R^{(\infty)},\mathscr{B}^{(\infty)},P^{(\infty)})$ 上的马氏序列 $\{\xi_1,\xi_2,\cdots\}$ 存在, 使得 $\{\xi_n\in B_n\}$ 在 ξ_{n-1} 之下关于 $P^{(\infty)}$ 的条件概率

$$P(\xi_n\in B_n|\xi_{n-1}=x_{n-1})\overset{\text{a.e.}}{=\!=}P_n(x_{n-1},B_n),B_n\in\mathscr{B}^{(1)},x_{n-1}\in R^{(1)},n=2,3,\cdots$$

而

$$P^{(\infty)}(\xi_1\in B_1)=P_1(B_1).$$

证明留作习题.

5.5.4 混合条件分布的存在及正则化定理 (可数维情形)

在这一小节, 我们将 §5.2 的定理 5 及 6 推广到 ξ_T 是随机变量序列的情形. 即

定理 9 设 $\xi_T=\{\xi_1,\xi_2,\cdots\}$ 是 $(\Omega,\mathscr{A},P)$ 上的随机变量序列,$\mathscr{C}$ 为 $\mathscr{A}$ 的任一子 σ-代数, 则 ξ_T 在 $\mathscr{C}$ 之下的混合条件分布存在. 若 ξ_T 的值域是 R^T 中的 Borel 集, 则 ξ_T 在 $\mathscr{C}$ 之下的条件分布存在.

证 显然定理的后一部分是定理前一部分及定理 4 的直接结果, 故只需证定理的前一部分.

当 ξ_T 是复随机函数时, 按定理 5 后面的注可以化作实随机函数的情形. 因而我们只对 ξ_T 是实随机函数的情形证明本定理.

考虑任意选定的一族在 $\mathscr{C}$ 之下关于 P 的条件概率 $P(A|\mathscr{C})(\omega),\omega\in\Omega,A\in\sigma(\xi_T)$. 我们先来证明, 可以找到自变数取有理数 $1,\cdots,r_n,\cdots$ 的函数序列

$$F^{\mathscr{C}}(\omega,r_1),\cdots,F^{\mathscr{C}}(\omega,r_1,\cdots,r_n)\cdots \tag{18}$$

具有下述性质:

$$\begin{cases}\Delta_{rr'}F^{\mathscr{C}}(\omega,\cdot)\geqslant 0, r\widehat{=}(r_1,\cdots,r_n)\geqslant(r_1',\cdots,r_n')\widehat{=}r',\\ \lim\limits_{r'\uparrow r}F^{\mathscr{C}}(\omega,r')=F^{\mathscr{C}}(\omega,r),\\ \lim\limits_{r_s\to-\infty}F^{\mathscr{C}}(\omega,r_1,\cdots,r_n)=0, 1\leqslant s\leqslant n,\\ \lim\limits_{\substack{r_k\to+\infty\\ k=1,\cdots,n}}F^{\mathscr{C}}(\omega,r_1,\cdots,r_n)=1,\\ F^{\mathscr{C}}(\omega,r')\leqslant F^{\mathscr{C}}(\omega,r), r'\leqslant r.\end{cases} \tag{19}$$

且对任何 $r_1,\cdots,r_n$,

$$F^{\mathscr{C}}(\cdot,r_1,\cdots,r_n)=P(\{\xi_1<r_1,\cdots,\xi_n<r_n\}|\mathscr{C}),\text{a.e.}, \tag{20}$$

此外还有

$$\lim_{r_{n+1}\to\infty}F^{\mathscr{C}}(\omega,r_1,\cdots,r_n,r_{n+1})=F^{\mathscr{C}}(\omega,r_1,\cdots,r_n). \tag{21}$$

事实上, 由定理 5 的证明知, 对每一 n, 存在 $\mathscr{C}$ 中的零概率集 $N^{(n)}$ 使得函数

$$F^{\mathscr{C}}(\omega,r_1,\cdots,r_n)=P(\xi_1<r_1,\cdots,\xi_n<r_n|\mathscr{C})(\omega),\omega\notin N^{(n)}$$

满足 (19) 中的各项条件.

其次, 由条件概率的性质知, 对每一正整数 n 及任何 $r^{(n)}=(r_1,\cdots,r_n)$, 有 $\mathscr{C}$ 中的零概率集 $N_r^{(n)}$, 使得当 $m\to\infty$ 时, 对一切 $\omega\in N^c_{r(n)}$ 有

$$P(\xi_1<r_1,\cdots,\xi_{n+1}<m|\mathscr{C})(\omega)\uparrow P(\xi_1<r_1,\cdots,\xi_n<r_n|\mathscr{C})(\omega). \tag{22}$$

取

$$N=\left(\bigcup_n N^{(n)}\right)\cup\left(\bigcup_n\left(\bigcup_{r(n)}N_{r(n)}\right)\right).$$

令

$$F^{\mathscr{C}}(\omega,r_1,\cdots,r_n)=\begin{cases}P(\xi_1<r_1,\cdots,\xi_n<r_n|\mathscr{C})(\omega),\omega\in N^c,\\ P(\xi_1<r_1,\cdots,\xi_n<r_n|\mathscr{C})(\omega_0),\omega\in N^c,\end{cases} \tag{23}$$

其中 ω_0 为任意取定的 N^c 中的元素. 易知 N 为可数个零概率集的并, 因而 $P(N)=0$.

由 F 的性质及 (22),(23) 知 $F^{\mathscr{C}}$ 满足 (19)—(21).

今再对一切 $\omega\in\Omega, n\geqslant 1, x^{(n)}=(x_1,\cdots,x_n)\in R^{(n)}$ 定义

$$F^{\mathscr{C}}(\omega,x_1,\cdots,x_n)=\lim_{r^{(n)}\uparrow x^{(n)}}F^{\mathscr{C}}(\omega,r_1,\cdots,r_n),$$

则由定理 5 的证明知对每一 $\omega\in\Omega, F^{\mathscr{C}}(\omega,x_1,\cdots,x_n), n=1,2,\cdots$ 是概率分布函数序列. 且由它们唯一决定的概率 $P_n^{\mathscr{C}}(\omega,B^{(n)}), B^{(n)}\in\mathscr{B}^{(n)}$ 是 $(\xi_1,\cdots,\xi_n)$ 在 $\mathscr{C}$ 之下的混合条件分布.

另一方面由 (21) 易证对每一 $\omega\in\Omega$, 概率场 $(R^{(n)},\mathscr{B}^{(n)},P_n^{\mathscr{C}}(\omega,\cdot)), n=1,2,\cdots$ 是和谐的. 故由定理 8 的证明知, 对每一 $\omega\in\Omega$, 有唯一的 $\mathscr{B}^T$ 上的概率 $P^{\mathscr{C}}(\omega,B^T)$, $B^T\in\mathscr{B}^T$, 使得

$$P^{\mathscr{C}}\left(\omega,B^{(n)}\times\prod_{\kappa=n+1}^{\infty}R_\kappa\right)=P_n^{\mathscr{C}}(\omega,B^{(n)}). \tag{24}$$

且由混合条件分布的定义知

$$\begin{aligned}P^{\mathscr{C}}\left(\omega,B^{(n)}\times\prod_{k=n+1}^{\infty}R_k\right)&=P_n^{\mathscr{C}}(\omega,B^{(n)})\\&=P\left(\xi_T\in B^{(n)}\times\prod_{k=n+1}^{\infty}R_k|\mathscr{C}\right)(\omega)\end{aligned} \tag{25}$$

对几乎一切 ω 成立.

最后, 我们来证明 $P^{\mathscr{C}}(\omega,B^T),(\omega,B^T)\in\Omega\times\mathscr{B}^T$ 是 ξ_T 的混合条件分布. 为此只需证明对每一 $B^T\in\mathscr{B}^T, P^{\mathscr{C}}(\omega,B^T)$ 是 $\mathscr{C}$-可测函数且

$$P^{\mathscr{C}}(\omega,B^T)=P(\xi_T\in B^T|\mathscr{C})(\omega) \tag{26}$$

对几乎一切 ω 成立.

首先由 (25) 及 $P_n^{\mathscr{C}}(\omega,B^{(n)})$ 的性质知上述断言对 $B^T=B^{(n)}\times\prod\limits_{k=n+1}^{\infty}R_k, B^{(n)}\in\mathscr{B}^{(n)}$ 成立. 再用 λ-系方法易证 $P^{\mathscr{C}}(\omega,B^T)$ 是 ξ_T 的混合条件分布. □

定理 10 设 $T=\{1,2,\cdots\}$, 则对概率空间 $(R^T,\mathscr{B}^T,P),\mathscr{C}$ 为 $\mathscr{B}^T$ 的任一子 σ-代数来说, 在 $\mathscr{C}$ 之下关于 P 的正则条件概率存在.

简言之, 可数维样本空间中的条件概率永远可以正则化.

证 令 $x^T=(x_1,x_2,\cdots)$ 表 R^T 中的点, 定义

$$\xi_n(x^T)=x_n, n=1,2,\cdots,x^T\in R^T,$$

则由定理 9 知 $\xi_T=(\xi_1,\xi_2,\cdots)$ 在 $\mathscr{C}$ 之下关于 P 的条件分布 $P^{\mathscr{C}}(x^T,B),B\in\sigma(\xi_T)$ 存在, 但由 $\sigma(\xi_T)=\mathscr{B}^T$ 立知 $P^{\mathscr{C}}(x^T,B),(x^T,B)\in R^T\times\mathscr{B}^T$ 即为 $\mathscr{C}$ 之下关于 P 的正则条件概率. □

由定理 10 易证

推论 2　设 $\xi_T=\{\xi_1,\xi_2,\cdots\}$ 是 $(\Omega,\mathscr{A},P)$ 上的随机变量序列, 则 ξ_T 在其样本空间 $(R^T,\mathscr{B}^T,P_{\xi_T})$ 上的表现 $\tilde{\xi}_T$ 在 $\mathscr{B}^T$ 的任一子 σ-代数 $\mathscr{C}$ 之下关于 P_{ξ_T} 的条件分布存在.

推论 3　设 $T=\{1,2,\cdots\}$, 则 $(R^T,\mathscr{B}^T,P)$ 上的任一随机函数 ξ_T 在 $\mathscr{B}^T$ 的任一子 σ-代数 $\mathscr{C}$ 下的条件分布存在.

习题及补充

1. 设 ξ_T 是 $(\Omega,\mathscr{A},P)$ 上的随机函数, $\mathscr{C}$ 是 $\mathscr{A}$ 的子 σ-代数, 则

i) $P^{\mathscr{C}}(\omega,A),(\omega,A)\in\Omega\times\sigma(\xi_T)$ 是 ξ_T 在 $\mathscr{C}$ 之下 (关于 P) 的条件分布的充分必要条件是下述二条件同时成立:

1) $P^{\mathscr{C}}(\omega,A)$ 是 $(\Omega,\mathscr{C})$ 到 $(\Omega,\sigma(\xi))$ 的转移概率;

2) 对一切 $A\in\sigma(\xi_T),B\in\mathscr{B}$,

$$\int_B P^{\mathscr{C}}(\cdot,A)dP=P(A\cap B).$$

ii) $P^{\mathscr{C}}(\omega,B),(\omega,B)\in\Omega\times\mathscr{B}^T$ 是 ξ_T 在 $\mathscr{C}$ 之下 (关于 P) 的混合条件分布的充分必要条件是下述二条件同时成立:

1) $P^{\mathscr{C}}(\omega,B)$ 是 $(\Omega,\mathscr{C})$ 到 $(R^T,\mathscr{B}^T)$ 的转移概率;

2) 对一切 $B^T\in\mathscr{B}^T,C\in\mathscr{C}$,

$$\int_C P^{\mathscr{C}}(\cdot,B^T)dP=P(C\cap\{\xi_T\in B^T\}).$$

2. 设 $\xi_T=(\xi_1,\xi_2,\cdots)$ 是 $(\Omega,\mathscr{A},P)$ 上的随机变量序列, 则 ξ_T 的分布律 P_{ξ_T} 决定一族转移概率序列 $P_1(B_1),P_2(x_1,B_2),\cdots,P_n(x_1,\cdots,x_{n-1},B_n),\cdots$ 使得对任何 $n,B^{(n)}\in\mathscr{B}^{(n)}$,

$$\begin{aligned}P_{\xi_T}\left(B^{(n)}\times\prod_{k=n+1}^{\infty}R_k\right)&=\int_{R_1}\int_{R_2}\cdots\int_{R_n}\chi_{B^{(n)}}(x_1,\cdots,x_n)\\&\quad\times P_n(x_1,\cdots,x_{n-1},dx_n)\cdots P_2(x_1,dx_2)P_1(dx_1).\end{aligned}$$

3. 完成例 1 的证明.

4. 设 $(\xi_1,\cdots,\xi_n)$ 是 $(R^{(n)},\mathscr{B}^{(n)},P)$ 上的样本函数, 且具有 n 维正态分布 $N(0,D)$, 令

$$\eta(x)=xD^{-1}x',x\in R^{(n)}.$$

试求 $\xi=(\xi_1,\cdots,\xi_n)$ 在 $\sigma(\eta)$ 之下关于 P 的条件分布.

第 6 章　特征函数及其初步应用

这一章我们将引进在概率论中有重要应用的一个分析工具, 即特征函数. 用数学分析的术语来说, 就是 Fourier-Stieltjes 变式.

特征函数之所以成为一个重要工具, 其原因是:

1° 分布律与特征函数之间存在一一对应关系. 因此当求出了随机变量的特征函数, 便可知其分布律. 由特征函数的某些性质, 可以推知分布律的相应性质.

不仅如此, 在分布律的某种收敛意义下的极限分布与特征函数的极限函数之间也存在着对应关系. 因此由特征函数的极限函数有时可以推知极限分布律, 因而推知随机变量序列的极限分布.

2° 特征函数是一种有界连续函数, 比分布函数及分布律更易于应用分析工具.

3° 独立随机变量, 特别是独立随机变量和以及与之有关的问题在概率论的发展中具有重要的地位, 为要研究独立随机变量和, 就要求出它的分布函数. 由第 4 章 §4.2 定理 5 知, 独立随机变量和的分布律是各随机变量分布律的卷积, 计算起来是很复杂的. 而独立随机变量和的特征函数等于它的各被加项的特征函数的乘积, 计算和研究都很方便. 这就是为什么古典极限问题能在引进特征函数之后很快得到解决的原因.

为了应用, 我们不仅研究随机变量概率分布的 Fourier-Stieltjes 变式, 也要研究一般的有界 L–S 测度的 Fourier-Stietjes 变式, 以及它们的收敛性之间的关系.

§6.1　特征函数的定义及初等性质

6.1.1　随机变量的特征函数的定义和例

定义 1　设 $(\xi_1,\cdots,\xi_n)$ 是 n 维实随机变量, 则称

$$f_{(\xi_1,\cdots,\xi_n)}(t_1,\cdots,t_n)\hat{=}E\left(\exp\left\{i\sum_{k=1}^{n}t_k\xi_k\right\}\right),(t_1,\cdots,t_n)\in R^{(n)}\tag{1}$$

为 $(\xi_1,\cdots,\xi_n)$ **的特征函数.**

为了书写简单, 今后经常采用矩阵的符号. 以 t, x 及 ξ 分别表示一行 n 列矩阵 $(t_1,\cdots,t_n),(x_1,\cdots,x_n)$ 及 $(\xi_1,\cdots,\xi_n)$, 用 A' 表示矩阵 A 的转置矩阵. 因而 $t\cdot\xi'=\sum\limits_{k=1}^{n}t_k\xi_k, t\cdot x'=\sum\limits_{k=1}^{n}t_kx_k$, 其中 $t\cdot x'$ 表示矩阵相乘. 于是 $\xi=(\xi_1,\cdots,\xi_n)$ 的

特征函数可以简记作

$$f_\xi(t) = E(\exp\{it\cdot\xi'\}).$$

因为 $|\exp\{it\cdot\xi'\}| = 1$, 对 P 可积, 故 ξ 的特征函数在 $R^{(n)}$ 上有定义, 并且由第 3 章 §3.4 知

$$\begin{aligned} f_\xi(t) &= \int_\Omega e^{it\cdot\xi'}dP = \int_{R^{(n)}} e^{it\cdot x'}dP_\xi \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{it\cdot x'}dF_\xi(x_1,\cdots,x_n). \end{aligned} \tag{2}$$

因此也称 $f_\xi(t)$ 为 ξ 的**分布律** P_ξ **或分布函数** F_ξ **的特征函数或 Fourier-Stieltjes 变式**.

若 ξ 为具有分布密度 $p(x_1,\cdots,x_n)$ 的连续型随机变量时, 则由第 3 章 §3.5 知

$$f_\xi(t) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{it\cdot x'}p(x_1,\cdots,x_n)dx_1\cdots dx_n. \tag{3}$$

若 ξ 是具有分布列

$$p_{l_1,\cdots,l_n} = P(\xi_1 = x_{1l_1},\cdots,\xi_n = x_{nl_n}), l_k = 1,2,\cdots, k=1,\cdots,n$$

的离散型随机变量, 则由第 3 章 §3.5 知其特征函数

$$f_\xi(t) = \sum_{l_1,\cdots,l_n}\exp\left\{i\sum_{k=1}^{n}t_k x_{kl_k}\right\}p_{l_1,\cdots,l_n}. \tag{4}$$

例 1　设 ξ 按二项分布律分布, 即 $P(\xi=k) = \binom{n}{k}p^k q^{n-k}, k=0,1,\cdots,n$, 则容易算出 ξ 的特征函数

$$f_\xi(t) = (q+pe^{it})^n. \tag{5}$$

若 ξ 是按多项分布律分布的 r 维随机变量, 即

$$P(\xi_1=k_1,\cdots,\xi_r=k_r) = \frac{n!}{k_1!\cdots k_r!k_{r+1}!}p_1^{k_1}\cdots p_r^{k_r}p_{r+1}^{k_{r+1}},$$

其中 $k_1+\cdots+k_r \leqslant n, k_s \geqslant 0, s=1,2,\cdots,r, k_1+\cdots+k_{r+1}=n, p_1+\cdots+p_{r+1}=1, p_s>0, s=1,\cdots,r+1$. 则由多项式定理知 ξ 的特征函数

$$\begin{aligned} f_\xi(t) &= \sum_{\substack{0\leqslant k_s\leqslant n\\ s=1,\cdots,r+1\\ k_1+\cdots+k_{r+1}=n}} \frac{n!}{k_1!\cdots k_{r+1}}(p_1e^{it_1})^{k_1}\cdots(p_re^{it_r})^{k_r}p_{r+1}^{k_{r+1}} \\ &= (p_1e^{it_1}+\cdots+p_re^{it_r}+p_{r+1})^n. \end{aligned} \tag{6}$$

例 2 若 ξ 具有参数 λ 的 Poisson 分布, 则 ξ 的特征函数

$$f_\xi(t)=\exp\{\lambda(e^{it}-1)\}. \tag{7}$$

若 ξ 具有参数 $\lambda_1,\cdots,\lambda_n$ 的 n 维 Poisson 分布, 则它的特征函数

$$f_\xi(t)=\exp\left\{\sum_{k=1}^{n}\lambda_k(e^{it_k}-1)\right\}=\prod_{k=1}^{n}\exp\left\{\lambda_k(e^{it_k}-1)\right\}. \tag{8}$$

例 3 设 ξ_r 按负二项分布律分布, 即

$$P(\xi_r=r+k)=\binom{-r}{k}p^r(-q)^k, k=0,1,2,\cdots,$$

则 ξ_r 的特征函数

$$f_{\xi_r}(t)=\sum_{k=0}^{\infty}e^{it(r+k)}\binom{-r}{k}p^r(-q)^k=\left(\frac{pe^{it}}{1-qe^{it}}\right)^r, \tag{9}$$

其中最后一个等式可由 Taylor 级数验证.

例 4 若 ξ 在 $[a,b]$ 上均匀分布, 则 ξ 的特征函数

$$f_\xi(t)=\int_a^b e^{itx}\frac{1}{b-a}dx=\begin{cases}\dfrac{e^{ibt}-e^{iat}}{i(b-a)t}, t\neq 0,\\ 1, \quad t=0.\end{cases} \tag{10}$$

例 5 设 ξ 按 $N(0,1)$ 分布, 则由 (3) 知 ξ 的特征函数

$$f_\xi(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{itx-\frac{x^2}{2}}dx.$$

由于

$$e^{itx-\frac{x^2}{2}}=e^{-\frac{(x-it)^2}{2}}e^{\frac{(it)^2}{2}},$$

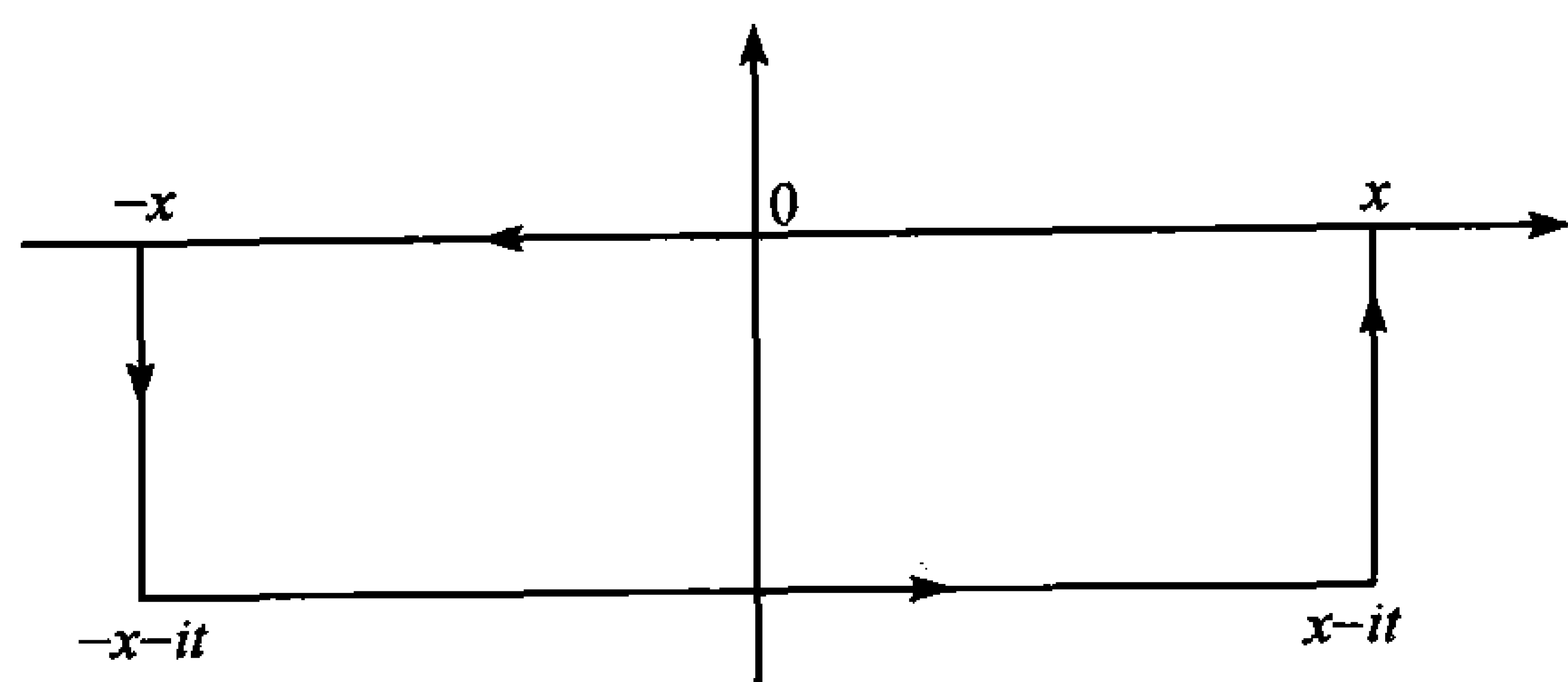
故

$$f_\xi(t)=e^{-\frac{t^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{(x-it)^2}{2}}dx. \tag{11}$$

为了算出上式右端的积分，将它了解为广义 R 积分而应用复变函数论中的 Cauchy 定理.

由于 $e^{-\frac{z^2}{2}}$ 是在整个复平面上的解析函数, 设 Γ 如图中标有箭头的线段组成的闭合曲线, 则

$$\begin{aligned}0&=\int_\Gamma e^{-\frac{z^2}{2}}dz\\ &=\int_{-x-it}^{x-it}e^{-\frac{z^2}{2}}dz+\int_{x-it}^{x}e^{-\frac{z^2}{2}}dz+\int_{x}^{-x}e^{-\frac{z^2}{2}}dz+\int_{-x}^{-x-it}e^{-\frac{z^2}{2}}dz\end{aligned}$$



即

$$\int_{-x-it}^{x-it} e^{-\frac{z^2}{2}}dz = \int_{-x}^{x} e^{-\frac{z^2}{2}}dz + \int_{x}^{x-it} e^{-\frac{z^2}{2}}dz + \int_{-x-it}^{-x} e^{-\frac{z^2}{2}}dz \tag{12}$$

对任何 $x>0$ 都成立. 但由复变函数积分的性质知

$$\left|\int_{x}^{x-it} e^{-\frac{z^2}{2}}dz\right| = \left|\int_0^t e^{-\frac{(x-is)^2}{2}}ds\right| \leqslant |t|e^{-\frac{x^2}{2}+\frac{t^2}{2}}.$$

当 $x\to\infty$ 时, 上述积分趋于零. 同理有

$$\lim_{x\to\infty}\int_{-x-it}^{-x} e^{-\frac{z^2}{2}}dz = 0.$$

故由 (12) 得

$$\begin{aligned}\int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}}dx &= \lim_{x\to\infty}\int_{-x-it}^{x-it} e^{-\frac{z^2}{2}}dz \\ &= \lim_{x\to\infty}\int_{-x}^{x} e^{-\frac{z^2}{2}}dz = \sqrt{2\pi}.\end{aligned}$$

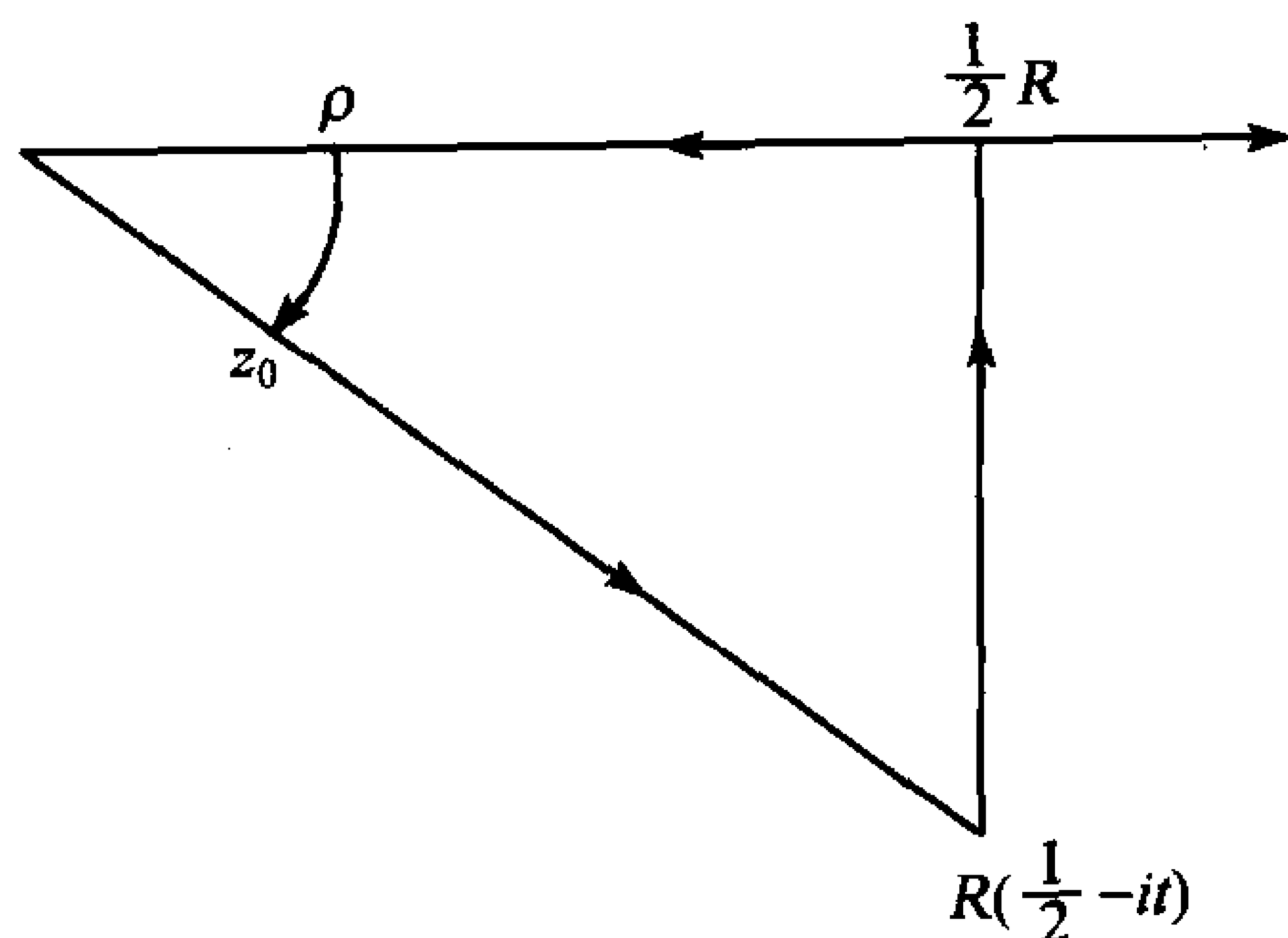
再由 (11) 式得

$$f_\xi(t) = e^{-\frac{t^2}{2}}.$$

例 6　设 ξ 按自由度为 n 的 χ^2-分布律分布, 则由 (3) 知 ξ 的特征函数

$$f_\xi(t) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\int_0^\infty e^{itx}e^{-\frac{x}{2}}x^{\frac{n}{2}-1}dx. \tag{13}$$

今往计算上式右端的积分. 由于 $e^{-z}z^{\frac{n}{2}-1}$ 在复平面上不包含零点和无穷远点的任何区域内解析. 设 Γ 如图中标有箭头的线段和圆弧组成的闭合曲线, 则根据 Cauchy 定理,



$$\int_\Gamma e^{-z}z^{(\frac{n}{2}-1)}dz=0,$$

即

$$\int_{z_0}^{R(\frac{1}{2}-it)} e^{-z}z^{\frac{n}{2}-1}dz=\int_\rho^{\frac{R}{2}} e^{-z}z^{\frac{n}{2}-1}dz+\int_{\frac{R}{2}}^{R(\frac{1}{2}-it)} e^{-z}z^{\frac{n}{2}-1}dz$$
$$+\int_{\substack{z_0\\|z|=\rho}}^{\rho} e^{-z}z^{\frac{n}{2}-1}dz. \tag{14}$$

由复变函数的积分性质易证当 $R\to\infty,\rho\to 0$ 时上式右端后两项积分的极限

$$\lim_{R\to\infty}\int_{\frac{R}{2}}^{R(\frac{1}{2}-it)} e^{-z}z^{\frac{n}{2}-1}dz=0, \tag{15}$$

$$\lim_{\rho\to 0}\int_{\substack{z_0\\|z|=\rho}}^{\rho} e^{-z}z^{\frac{n}{2}-1}dz=0. \tag{16}$$

其次设 $z_0=\left(\dfrac{1}{2}-it\right)y_0$, 则

$$\int_{z_0}^{R(\frac{1}{2}-it)} e^{-z}z^{\frac{n}{2}-1}dz=\left(\frac{1}{2}-it\right)^{\frac{n}{2}}\int_{y_0}^{R} e^{(it-\frac{1}{2})x}x^{\frac{n}{2}-1}dx. \tag{17}$$

故当 $R\to\infty,\rho\to 0$ (注意此时 $y_0\to 0$) 由 (14) − (17) 即得

$$\left(\frac{1}{2}-it\right)^{\frac{n}{2}}\int_0^{\infty} e^{(it-\frac{1}{2})x}x^{\frac{n}{2}-1}dx=\int_0^{\infty} e^{-x}x^{\frac{n}{2}-1}dx=\Gamma\left(\frac{n}{2}\right).$$

再由 (13) 知

$$f_\xi(t)=(1-2it)^{-\frac{n}{2}}.$$

用类似的方法可以求出 Γ- 分布的特征函数.

6.1.2　随机变量特征函数的初等性质

如无特别声明, 以下总用 $f_\xi(t)$(或 $f(t)$) 表示 $\xi=(\xi_1,\cdots,\xi_n)$ 的特征函数 $f_\xi(t_1,\cdots,t_n)$.

性质 1　$f_\xi(t)$ 具有下列简单性质:

a) $f_\xi(0)=1$;

b) $|f_\xi(t)|\leqslant 1, t\in R^{(n)}$;

c) $\overline{f_\xi(t)}=f_\xi(-t)$, 其中 $-t=(-t_1,\cdots,-t_n)$;

d) $f_{-\xi}(t)=\overline{f_\xi(t)}$.

证　a) 及 b) 的证明: 由积分性质知对任何 $t\in R^{(n)}$ 来说,

$$|f_\xi(t)|=\left|\int_\Omega e^{it\cdot\xi'}dp\right|\leqslant\int_\Omega|e^{it\cdot\xi'}|dP=\int_\Omega dP=f_\xi(0)=1.$$

c) 及 d) 的证明: 因为

$$\begin{aligned}\overline{f_\xi(t)}&=\int_\Omega\overline{e^{it\cdot\xi'}}dP=\int_\Omega e^{-it\cdot\xi'}dP\\&=\int_\Omega e^{i(-t)\cdot\xi'}dP=f_\xi(-t)\\&=\int_\Omega e^{it\cdot(-\xi)'}dP=f_{-\xi}(t).\end{aligned}$$

□

性质 2　对于任何 $1\leqslant i_1<\cdots<i_k\leqslant n, 1\leqslant k\leqslant n$ 来说, $(\xi_{i_1},\cdots,\xi_{i_k})$ 的特征函数

$$f(t_{i_1,\cdots,t_{i_k}})=f_\xi(0,\cdots,0,t_{i_1},0\cdots,0,t_{i_k},0,\cdots,0).$$

性质 3　设 $A=(a_{kl})$ 是一 $m\times n$ 矩阵, $b=(b_1,\cdots,b_m)$, 随机变量 $\eta=\xi A'+b$ 的特征函数

$$f_\eta(u_1,\cdots,u_m)=e^{iu\cdot b'}f_\xi(u\cdot A).\tag{18}$$

特别, $\xi_1+\cdots+\xi_n$ 的特征函数

$$f(t)=f_\xi(t,\cdots,t).$$

当 ξ 为一维实随机变量时, $a\xi+b$ (其中 $a\neq 0, a, b$ 为实数) 的特征函数

$$f_{a\xi+b}(t)=e^{itb}f_\xi(at).$$

证　由定义知

$$f_\eta(u)=E(e^{iu\cdot\eta'})=E(\exp\{iu\cdot A.\xi'+iu\cdot b'\})$$

$$= e^{iu\cdot b'}E(\exp\{i(u\cdot A)\cdot\xi'\}) = e^{iu\cdot b'}f_\xi(u\cdot A).$$ □

由性质 3 的证明, 对于 ξ 的任一组实值 Borel 可测函数 $\eta=(\eta_1,\cdots,\eta_m)=(g_1(\xi_1,\cdots,\xi_n),\cdots,g_m(\xi_1,\cdots,\xi_n))$ 的特征函数可如下求得:

$$f_\eta(u_1,\cdots,u_m)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\exp\left\{i\sum_{l=1}^m u_l g_l(x_1,\cdots,x_n)\right\}\times dF(x_1,\cdots,x_n).$$

性质 4 若 $\xi^{(k)}=(\xi_1^{(k)},\cdots,\xi_{m_k}^{(k)}),k=1,\cdots,n$ 是独立的随机变量, $\xi=(\xi_1^{(1)},\cdots,\xi_{m_1}^{(1)},\cdots,\xi_1^{(n)},\cdots,\xi_{m_n}^{(n)})$, 则

$$f_\xi(t)=\prod_{k=1}^n f_{\xi^{(k)}}(t^{(k)}),\tag{19}$$

其中 $t=(t_1^{(1)},\cdots,t_{m_n}^{(n)}),t^{(k)}=(t_1^{(k)},\cdots,t_{m_k}^{(k)}),k=1,\cdots,n$. 特别, 若 $m_1=\cdots=m_n=m$, 则独立随机变量 $\xi^{(k)},k=1,\cdots,n$ 的和 $\xi^{(1)}+\cdots+\xi^{(n)}$ 的特征函数

$$f(t)=\prod_{k=1}^n f_{\xi^{(k)}}(t),t=(t_1,\cdots,t_m).$$

证 由第 2 章 §2.4 推论 3 知 $\exp\{e^{it^{(k)}\cdot\xi^{(k)'}}\},k=1,\cdots,n$ 是独立的, 故由数学期望的乘法定理知

$$f_\xi(t)=E(\exp\{it\cdot\xi'\})=E\left(\exp\left\{i\sum_{k=1}^n t^{(k)}\cdot\xi^{(k)'}\right\}\right)=E\left(\prod_{k=1}^n\exp\{it^{(k)}\cdot\xi^{(k)'}\}\right)$$
$$=\prod_{k=1}^n E(\exp\{it^{(k)}\cdot\xi^{(k)'}\})=\prod_{k=1}^n f_{\xi^{(k)}}(t^{(k)}).$$

此即第一部分的结论. 当 $m_1=\cdots=m_n=m$ 时, 在上式中令 $t^{(1)}=\cdots=t^{(n)}=\bar t=(t_1,\cdots,t_m)$, 则得

$$f_{\xi^{(1)}+\cdots+\xi^{(n)}}(\bar t)=E\left(\exp\left\{i\bar t\cdot\left(\sum_{k=1}^n\cdot\xi^{(k)}\right)'\right\}\right)$$
$$=\prod_{k=1}^n E(\exp\{i\bar t\cdot\xi^{(k)'}\})=\prod_{k=1}^n f_{\xi^{(k)}}(\bar t).$$

此即第二部分结论.

现在我们应用以上各性质计算一些随机变量的特征函数.

例 7 设 ξ 是按 $N(a,\sigma^2)$ 分布的, 则易知 $\eta=\dfrac{\xi-a}{\sigma}$ 仍然是正态分布, 且

$E\eta=0, D\eta=1$. 故 η 按 $N(0,1)$ 分布, 于是由性质 3 及例 5 知

$$f_\xi(t)=f_{\sigma\eta+a}(t)=e^{ita}e^{-\frac{(t\sigma)^2}{2}}=e^{iat-\frac{\sigma^2t^2}{2}}.$$

例 8　设 $\xi=(\xi_1,\cdots,\xi_n)$ 是按 $N(m,D)$ 分布的, 则由第 4 章 §4.2 例 4 知, 有一正交方阵 C 存在使得

$$\eta=\xi\cdot C'-mC'$$

的分布密度为

$$\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi}\sigma_k}e^{-\frac{x^2}{2\sigma_k^2}},$$

因而 $\eta_1,\cdots,\eta_n$ 独立. 令

$$\Lambda=CDC'=\begin{pmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_n^2\end{pmatrix}.$$

于是由性质 4 及例 7 知

$$\begin{aligned}f_\eta(t)&=\prod_{k=1}^{n}f_{\eta_k}(t_k)=\exp\left\{-\frac{1}{2}\sum_{k=1}^{n}\sigma_k^2t_k^2\right\}\\&=\exp\left\{-\frac{1}{2}t\Lambda t'\right\}=\exp\left\{-\frac{1}{2}tCD(tC)'\right\}.\end{aligned}$$

故由性质 3 及 $\xi=(\eta+mC')C=\eta C+m$ 知

$$\begin{aligned}f_\xi(u)&=e^{iu\cdot m'}f_\eta(u\cdot C')\\&=e^{iu\cdot m'}\exp\left\{-\frac{1}{2}uC'CD(uC'C)'\right\}\\&=\exp\left\{iu\cdot m'-\frac{1}{2}uDu'\right\}.\end{aligned}$$

例 9　若 $\xi=(\xi_1,\cdots,\xi_n)$ 按 $N(m,D)$ 分布, 则由例 8 及性质 3 知 $\xi_1+\cdots+\xi_n$ 的特征函数

$$f_{\xi_1+\cdots+\xi_n}(t)=\exp\left\{it\sum_{k=1}^{n}m_k-\frac{t^2}{2}\sum_{k,l=1}^{n}d_{kl}\right\},$$

其中 $D=(d_{kl})$. 若 $\xi_1,\cdots,\xi_n$ 独立, 且 ξ_k 按 $N(m_k,\sigma_k^2)$ 分布, 则

$$D=\begin{pmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_n^2\end{pmatrix}, m=(m_1,\cdots,m_n).$$

因而

$$f_{\xi_1+\cdots+\xi_n}(t)=\exp\left\{it\sum_{k=1}^n m_k-\frac{t^2}{2}\sum_{k=1}^n\sigma_k^2\right\}.$$

在下一节证明了分布函数由特征函数唯一确定后, 我们就可得到 $\xi_1+\cdots+\xi_n$ 在前一种情形下是按 $N\left(\sum_{k=1}^n m_k,\sum_{k,l=1}^n d_{kl}\right)$ 分布的, 在后一种情形下是按 $N\left(\sum_{k=1}^n m_k,\sum_{k=1}^n\sigma_k^2\right)$ 分布的.

例 10 设 $\xi=(\xi_1,\cdots,\xi_n)$ 是按 $N(m,D)$ 分布的, $k,i_1,\cdots,i_k$ 满足条件:$1\leqslant k\leqslant n, 1\leqslant i_1<\cdots<i_k\leqslant n$, 则由性质 2 及例 8 知 $(\xi_{i_1},\cdots,\xi_{i_k})$ 的特征函数

$$f(u_1,\cdots,u_k)=\exp\left\{i\sum_{l=1}^k u_l m_{i_l}-\frac{1}{2}\sum_{l,m=1}^k d_{i_l i_m}u_l u_m\right\}.$$

在下一节证明了分布函数由特征函数唯一确定后, 我们就可断言 $(\xi_{i_1},\cdots,\xi_{i_k})$ 是按 $N(\bar{m},\bar{D})$ 分布的, 其中

$$\bar{m}=(m_{i_1},\cdots,m_{i_k}),$$

而 $\bar{D}$ 为方阵 D 在第 $i_1,\cdots,i_k$ 各行及各列交叉处的元素所成的方阵. 这个结论充分说明了特征函数对分布的计算的有效性. 读者不难通过亲身试算得知应用累次积分计算得出这个结论会遭遇到多大的困难.

例 11 设 χ_k^2 是自由度为 n_k 的 χ^2-变量, $k=1,\cdots,r$, 且相互独立, 则由例 6 及性质 4 知 $\sum_{k=1}^r\chi_k^2$ 的特征函数是

$$\prod_{k=1}^r(1-2it)^{-\frac{n_k}{2}}=(1-2it)^{-\frac{1}{2}(n_1+\cdots+n_r)}$$

由下节特征函数唯一决定分布函数的定理知 $\sum_{k=1}\chi_k^2$ 是自由度为 $n_1+\cdots+n_r$ 的 χ^2-变量.

性质 5(增量不等式)　设 $f(t)$ 是 ξ 的特征函数, 则对任何 $t=(t_1,\cdots,t_n)\in R^{(n)}$及$h=(h_1,\cdots,h_n)\in R^{(n)}$,

$$|f(t)-f(t+h)|^2\leqslant 2f(0)\{f(0)-Rf(h)\}, \tag{20}$$

其中 Rz 表示 z 的实部, 且 $f(t)$ 在 $R^{(n)}$ 上一致连续.

证　由定义知

$$\begin{aligned}f(t)-f(t+h)&=\int[e^{it\cdot\xi'}-e^{i(t+h)\cdot\xi'}]dP\\&=\int e^{it\cdot\xi'}[1-e^{ih\cdot\xi'}]dP.\end{aligned}$$

应用 Schwarz 不等式知

$$\begin{aligned}&|f(t)-f(t+h)|^2\leqslant\int|e^{it\cdot\xi'}|^2dP\cdot\int|1-e^{ih\cdot\xi'}|^2dP\\&=f(0)\cdot\int[(1-\cos h\cdot\xi')^2+(\sin h\cdot\xi')^2]dP\\&=2f(0)\int[1-\cos h\cdot\xi']dP=2f(0)(f(0)-Rf(h)).\end{aligned}$$

由控制收敛定理知

$$\lim_{h\to0}\int[1-\cos h\cdot\xi']dP=\int\lim_{h\to0}[1-\cos h\cdot\xi']dP=0.$$

故 $f(t)$ 在零点连续, 再由 (20) 知 $f(t)$ 在 $R^{(n)}$ 上一致连续.

性质 6　设 ξ 是一维实随机变量, $f(t)$ 是它的特征函数, $f^{(k)}(t)$ 表示 $f(t)$ 的 k 阶导数.

i) 若 $f^{(2n)}(0)$ 存在且有限, 则对任一满足 $0\leqslant r\leqslant 2n$ 的 r 来说, 有 $\beta_r=E|\xi|^r<\infty$.

ii) 若存在一 $\delta\geqslant0,\xi$ 的 $n+\delta$ 绝对矩 $\beta_{n+\delta}=E|\xi|^{n+\delta}$ 有限, 则对一切 $k\leqslant n$ 恒有

$$f^{(k)}(t)=i^k\int_{-\infty}^{\infty}x^ke^{itx}dF_\xi(x),t\in R^{(1)}. \tag{21}$$

因而

$$\alpha_k=E\xi^k=i^{-k}f^{(k)}(0), \tag{22}$$

且 $f^{(k)}(t)$ 在 $R^{(1)}$ 上一致连续并以 β_k 为界, 此外还有

$$f(t)=\sum_{k=0}^{n-1}\alpha_k\frac{(it)^k}{k!}+\rho_n(t),t\in R^{(1)}, \tag{23}$$

其中

$$\rho_n(t) = t^n \int_0^1 \frac{(1-u)^{n-1}}{(n-1)!} f^{(n)}(tu)du$$
$$= \alpha_n \frac{(it)^n}{n!} + o(|t|^n)$$
$$= \theta\beta_n \frac{|t|^n}{n!}, |\theta| \leqslant 1. \tag{24}$$

若 $0 < \delta \leqslant 1$, 则

$$\rho_n(t) = \alpha_n \frac{(it)^n}{n!} + 2^{1-\delta}\theta'\beta_{n+\delta}\frac{|t|^{n+\delta}}{(1+\delta)(2+\delta)\cdots(n+\delta)}, \quad |\theta'| \leqslant 1. \tag{25}$$

证 i) 先证明一个分析求导公式: 若 $f^{(n)}(0)$ 存在, 则

$$f^{(n)}(0) = \lim_{h\to 0}\frac{1}{(2h)^n}\sum_{k=0}^{n}\binom{n}{k}(-1)^k f((n-2k)h). \tag{26}$$

事实上, 利用 Taylor 公式我们有

$$f(x) = \sum_{j=0}^{n} f^{(j)}(0)\frac{x^j}{j!} + o(|x|^n).$$

于是

$$\frac{1}{(2h)^n}\sum_{k=0}^{n}\binom{n}{k}(-1)^k f((n-2k)h)$$
$$= \frac{1}{(2h)^n}\sum_{k=0}^{n}\binom{n}{k}(-1)^k\left[\sum_{j=0}^{n} f^{(j)}(0)\frac{(n-2k)^j h^j}{j!} + o(h^n)\right]$$
$$= \frac{1}{(2h)^n}\sum_{j=0}^{n} f^{(j)}(0)\frac{h^j}{j!}\sum_{k=0}^{n}\binom{n}{k}(-1)^k(n-2k)^j + o(1).$$

利用恒等式

$$\sum_{k=0}^{n}\binom{n}{k}(-1)^k e^{(n-2k)t} = (e^t - e^{-t})^n$$

两边求 j 阶导数 ($j \leqslant n$), 再令 $t = 0$ 可证

$$\sum_{k=0}^{n}\binom{n}{k}(-1)^k(n-2k)^j = \begin{cases} 0, & 0 \leqslant j < n, \\ 2^n \cdot n!, & j = n. \end{cases}$$

即得 (26) 式.

再应用 (26) 式求 ξ 的特征函数在零点的导数. 若 $f^{(m)}(0)$ 存在, 则

$$f^{(m)}(0) = \lim_{h\to 0}\frac{1}{(2h)^m}\sum_{k=0}^{m}\binom{m}{k}(-1)^k f((m-2k)h)$$

$$= \lim_{h\to 0}\frac{1}{(2h)^m}\int_{-\infty}^{\infty}\sum_{k=0}^{m}\binom{m}{k}(-1)^k e^{i(m-2k)hx}dF(x)$$

$$= \lim_{h\to 0}\frac{1}{(2h)^m}\int_{-\infty}^{\infty}(e^{ihx}-e^{-ihx})^m dF(x)$$

$$= \lim_{h\to 0} i^m\int_{-\infty}^{\infty}\left(\frac{\sin hx}{hx}\right)^m x^m dF(x).$$

若 $f^{(2n)}(0)$ 存在, 有限, 则

$$\infty > |f^{(2n)}(0)| = \lim_{h\to 0}\int\left(\frac{\sin hx}{hx}\right)^{2n} x^{2n}dF(x)$$

$$\geqslant \int \varliminf_{h\to 0}\left(\frac{\sin hx}{hx}\right)^{2n} x^{2n}dF(x)$$

$$= \int x^{2n}dF(x) = \beta_{2n}.$$

式中不等号成立是由于 Fatou-Lebesgue 定理. 这就证明了 ξ 的 $2n$ 阶矩 β_{2n} 存在有限. 再由第 3 章 3.6.1 不等式 (6) 知, 对一切 $0\leqslant r\leqslant 2n, \beta_r$ 有限.

ii) 由于当 $r\leqslant n+\delta$ 时

$$[E|\xi|^r]^{\frac{1}{r}} \leqslant [E|\xi|^{n+\delta}]^{\frac{1}{n+\delta}} < \infty,$$

故对一切 $k\leqslant n$, 有

$$\int_{-\infty}^{\infty}|x|^k dF(x) = E|\xi|^k < \infty.$$

而

$$f(t) = \int_{-\infty}^{\infty} e^{itx}dF(x).$$

今往证: 当 $E|\xi|^k<\infty$ 时 $f^{(k)}(t)$ 存在且

$$f^{(k)}(t) = i^k\int x^k e^{itx}dF(x).$$

事实上, 当 $k=1$ 时, 由第 3 章 3.6.3 推论 6 及

$$\left|\frac{de^{itx}}{dt}\right| = |ixe^{itx}| = |x|$$

而 x 关于 P_ξ 可积, $\int e^{itx}dF_\xi(x)$ 有限知:

$$f'(t)=\int ixe^{itx}dF(x).$$

若 (21) 式对 $k-1$ 成立, $f^{(k-1)}(t)=i^{k-1}\int x^{k-1}e^{itx}dF_\xi(x)$ 存在, 有限, 则由

$$\left|\frac{d(i^{k-1}x^{k-1}e^{itx})}{dt}\right|=|x^ke^{itx}|=|x^k|$$

关于 P_ξ 可积知

$$f^{(k)}(t)=i^k\int x^ke^{itx}dF_\xi(x),$$

且

$$|f^{(k)}(t)-f^{(k)}(t+h)|\leqslant\int|1-e^{ihx}||x|^kdF_\xi(x). \tag{27}$$

因而 $f^{(k)}(t)$ 在 $R^{(1)}$ 上一致连续且以 β_k 为界.

在 (21) 式中令 $t=0$ 即得 (22) 式.

iii) 往证 (23) 式. 我们先证明公式 (23) 具有 (24) 中第一尾式. 用分部积分法

$$\begin{aligned}
&t^n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}f^{(n)}(tu)du\\
=&t^{n-1}f^{(n-1)}(tu)\frac{(1-u)^{n-1}}{(n-1)!}\bigg|_0^1+t^{n-1}\int_0^1\frac{(1-u)^{n-2}}{(n-2)!}f^{(n-1)}(tu)du\\
=&-\frac{t^{n-1}}{(n-1)!}f^{(n-1)}(0)+t^{n-1}\int_0^1\frac{(1-u)^{n-2}}{(n-2)!}f^{(n-1)}(tu)du\\
=&\cdots=-\sum_{k=0}^{n-1}\frac{t^k}{k!}f^{(k)}(0)+f(t).
\end{aligned}$$

而由 (22) 式知 $f^{(k)}(0)=i^k\alpha_k$, 将此式代入上式即得 (23) 具有 (24) 中第一尾式.

其次由 (22) 及 (27) 知

$$\begin{aligned}
\left|t^n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}f^{(n)}(tu)du-\alpha_n\frac{(it)^n}{n!}\right|&=\left|t^n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}[f^{(n)}(tu)-f^{(n)}(0)]du\right|\\
&\leqslant|t|^n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}|f^{(n)}(tu)-f^{(n)}(0)|du\\
&\leqslant\varepsilon|t|^n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}du=\varepsilon\frac{|t|^n}{n!},
\end{aligned}$$

其中 $\varepsilon>0$ 可以取得任意小 (只要 t 足够小), 因而 (23) 具有 (24) 中第二尾式.

第三, 我们证明 (23) 具有 (24) 中第三尾式. 因为 $f^{(n)}(t)$ 以 β_n 为界, 故

$$\left|t^n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}f^{(n)}(tu)du\right|\leqslant|t|^n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}|f^{(n)}(tu)|du$$
$$\leqslant|t|^n\beta_n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}du=\beta_n\frac{|t|^n}{n!}.$$

故

$$\rho_n(t)=\theta\beta_n\frac{|t|^n}{n!},|\theta|\leqslant 1.$$

iv) 最后来证明: 当 $0<\delta\leqslant 1$ 时, (25) 式成立. 由 (27) 式及 (24) 的第一尾式知

$$\left|\rho_n(t)-\alpha_n\frac{(it)^n}{n!}\right|=\left|t^n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}f^{(n)}(tu)du-\alpha_n\frac{(it)^n}{n!}\right|$$
$$\leqslant|t|^n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}|f^{(n)}(tu)-f^{(n)}(0)|du$$
$$\leqslant|t|^n\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}\int_{-\infty}^{\infty}|1-e^{iutx}||x|^ndF(x)du.$$

但是容易证明 $|1-e^{iuxt}|\leqslant 2\left|\frac{tux}{2}\right|^\delta$ $\left(\text{因为若 }\left|\frac{tux}{2}\right|<1,\text{ 则 }|1-e^{itux}|=[2(1-\cos tux)]^{\frac{1}{2}}\leqslant|tux|\leqslant 2\left|\frac{tux}{2}\right|^\delta,\text{ 若 }\left|\frac{tux}{2}\right|\geqslant 1,\text{ 则 }|1-e^{itux}|\leqslant 2\leqslant 2\left|\frac{tux}{2}\right|^\delta\right)$, 故由分部积分法即得

$$\left|\rho_n(t)-\alpha_n\frac{(it)^n}{n!}\right|\leqslant|t|^{n+\delta}\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}u^\delta\int_{-\infty}^{\infty}2^{1-\delta}|x|^{n+\delta}dF(x)du$$
$$=2^{1-\delta}\beta_{n+\delta}|t|^{n+\delta}\int_0^1\frac{(1-u)^{n-1}}{(n-1)!}u^\delta du$$
$$=2^{1-\delta}\beta_{n+\delta}\frac{|t|^{n+\delta}}{(1+\delta)\cdots(n+\delta)}.$$

因此 (25) 式获证. □

推论　若 ξ 的一切阶矩存在, 则

$$f_\xi(t)=\sum_{n=0}^{\infty}\alpha_n\frac{(it)^n}{n!}\tag{28}$$

在上式右端级数收敛的区间中成立. 特别若 α_n 有界, 则上式在 $R^{(1)}$ 上成立.

证 由于 ξ 的一切阶矩存在, 由性质 6 知

$$f(t)=\sum_{k=0}^{n-1}\alpha_k\frac{(it)^k}{k!}+\rho_n(t),$$

而 $\rho_n(t)=\theta\beta_n\frac{|t|^n}{n!}, |\theta|\leqslant 1$. 我们只要证明在 (28) 式右端级数的收敛区间上 $\rho_n(t)\to 0$ 即可. 事实上, 若 (28) 右端收敛, 必有 $\alpha_n\frac{(it)^n}{n!}\to 0$, 而当 $n=2k$ 时 (k 为正整数)$\beta_n=\alpha_n$, 因而 $\rho_{2k}(t)=\theta\beta_{2k}\frac{|t|^{2k}}{(2k)!}=\theta\alpha_{2k}\frac{|t|^{2k}}{(2k)!}\to 0, (k\to\infty)$. 应用 Schwarz 不等式得

$$\begin{aligned}(\beta_{2k+1})^2&=\left[\int_{-\infty}^{\infty}|u|^{2k+1}dF(u)\right]^2\leqslant\int|u|^{2k}dF(u)\int|u|^{2k+2}dF(u)\\&=\beta_{2k}\cdot\beta_{2(k+1)},\end{aligned}$$

因而当 $k\to\infty$ 时,

$$\begin{aligned}|\rho_{2k+1}(t)|^2&\leqslant\beta_{2k}\cdot\beta_{2(k+1)}\frac{(|t|^{2k+1})^2}{[(2k+1)!]^2}\\&=\frac{\beta_{2k}|t|^{2k}}{(2k)!}\frac{\beta_{2(k+1)}|t|^{2(k+1)}}{(2k+2)!}\frac{2k+2}{2k+1}\to 0,\end{aligned}$$

即 (28) 式在级数收敛的区间中成立. □

利用性质 6 及其推论, 使用微分法可求得随机变量的各阶矩, 它比第 3 章中由积分求各阶矩的方法在计算上要简单得多.

例 12 求负二项分布 ξ_r 的矩. 由例 3 知它的特征函数

$$\begin{aligned}f(t)&=\left(\frac{pe^{it}}{1-qe^{it}}\right)^r=p^r(e^{-it}-q)^{-r},\\f'(t)&=irp^re^{-it}(e^{-it}-q)^{-(r+1)},\\f''(t)&=i^2rp^re^{-it}(e^{-it}-q)^{-(r+2)}(q+re^{-it}).\end{aligned}$$

由 $f''(0)$ 存在推知 β_2 存在, 因而 α_1,α_2 即 $E\xi, D\xi$ 都存在, 并且

$$\begin{aligned}E\xi&=\alpha_1=\frac{f'(0)}{i}=rp^r(1-q)^{-(r+1)}=\frac{r}{p},\\E\xi^2&=\alpha_2=\frac{f''(0)}{i^2}=rp^r(1-q)^{-(r+2)}(q+r)=\frac{r(r+q)}{p^2},\end{aligned}$$

$$D\xi = E\xi^2 - (E\xi)^2 = \frac{rq}{p^2}.$$

例 13　求 $N(0,\sigma^2)$ 分布的各阶矩.

由于 $N(0,\sigma^2)$ 的特征函数为

$$f(t) = e^{-\frac{\sigma^2 t^2}{2}} = \sum_{k=0}^{\infty}(-1)^k t^{2k}\frac{\left(\frac{\sigma^2}{2}\right)^k}{k!},$$

而 $f(t)$ 的幂级数展开式在 $R^{(1)}$ 上收敛, 因而 $f(t)$ 具有一切阶导数, 由性质 6 知 $N(0,\sigma^2)$ 具有一切阶矩. 再由

$$\left.\begin{aligned} f^{(2k)}(0) &= (-1)^k\frac{(2k)!\sigma^{2k}}{2^k\cdot k!}, \\ f^{(2k-1)}(0) &= 0, \end{aligned}\right. \qquad k = 1,2,\cdots$$

知

$$\left.\begin{aligned} \alpha_{2k} &= \frac{f^{(2k)}(0)}{i^{2k}} = \frac{(2k)!\sigma^{2k}}{2^k\cdot k!}, \\ \alpha_{2k-1} &= 0, \end{aligned}\right. \qquad k = 1,2,\cdots.$$

若 ξ 是按 $N(a,\sigma^2)$ 分布的, 则 $\xi - a$ 是按 $N(0,\sigma^2)$ 分布的, 于是得 ξ 的中心矩为

$$E(\xi-a)^n = \begin{cases} 0, & n = 2k-1, \\ \dfrac{(2k)!\sigma^{2k}}{2^k\cdot k!}, & n = 2k, \end{cases} \qquad k = 1,2,\cdots.$$

性质 7　若对一切满足条件 $r_1 + \cdots + r_n = m$ (m 为正整数) 的非负整数 $r_1,\cdots,r_n$ 来说, $E|\xi_1^{r_1}\cdots\xi_n^{r_n}|$ 有限, 则对任何满足 $k_1+\cdots+k_n \leqslant m$ 的非负整数 $k_1,\cdots,k_n, E|\xi_1^{k_1}\cdots,\xi_n^{k_n}| < \infty$ 且

$$\frac{\partial^{k_1+\cdots+k_n}}{\partial t_1^{k_1}\cdots\partial t_n^{k_n}} f_\xi(t_1,\cdots,t_n) = i^{k_1+\cdots+k_n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} e^{it\cdot x'} x_1^{k_1}\cdots x_n^{k_n} dF_\xi(x_1,\cdots,x_n). \tag{29}$$

因而

$$\begin{aligned} \alpha_{k_1,\cdots,k_n} &= E\xi_1^{k_1}\cdots\xi_n^{k_n} \\ &= i^{-(k_1+\cdots+k_n)}\frac{\partial^{k_1+\cdots+k_n}}{\partial t_1^{k_1}\cdots\partial t_n^{k_n}} f_\xi(t_1,\cdots,t_n)|_{t=0} \end{aligned} \tag{30}$$

且

$$f_\xi(t)=\sum_{\substack{k_1+\cdots+k_n\leqslant m-1\\ k_1,\cdots,k_n\geqslant 0}}\alpha_{k_1\cdots k_n}\frac{(it_1)^{k_1}}{k_1!}\cdots\frac{(it_n)^{k_n}}{k_n!}+\rho_m(t), \tag{31}$$

其中

$$\begin{aligned}\rho_m(t)&=\sum_{\substack{k_1+\cdots+k_n=m\\ k_1,\cdots,k_n\geqslant 0}}\alpha_{k_1\cdots k_n}\frac{(it_1)^{k_1}}{k_1!}\cdots\frac{(it_n)^{k_n}}{k_n!}+o(|t_1^2+\cdots+t_n^2|^{\frac{m}{2}})\\ &=\theta\sum_{\substack{k_1+\cdots+k_n=m\\ k_1,\cdots,k_n\geqslant 0}}\beta_{k_1},\cdots,{}_{k_n}\frac{|t_1|^{k_1}\cdots|t_n|^{k_n}}{k_1!\cdots k_n!}.\end{aligned} \tag{32}$$

证 i) 先证对任何满足 $k_1+\cdots+k_n<m$ 的非负整数 $k_1,\cdots,k_n$, $E|\xi_1^{k_1}\cdots\xi_n^{k_n}|<\infty$. 事实上,

$$\begin{aligned}E|\xi_1^{k_1}\cdots\xi_n^{k_n}|&=\int_{|\xi_1|<1}|\xi_1^{k_1}\cdots\xi_n^{k_n}|dP+\int_{|\xi_1|\geqslant 1}|\xi_1^{k_1}\cdots\xi_n^{k_n}|dP\\ &\leqslant\int|\xi_2^{k_2}\cdots\xi_n^{k_n}|dP+\int|\xi_1^{m-(k_2+\cdots+k_n)}\xi_2^{k_2}\cdots\xi_n^{k_n}|dP\\ &\leqslant\cdots\\ &\leqslant\sum_{j=1}^{n}\int|\xi_j^{m-(k_{j+1}+\cdots+k_n)}\xi_{j+1}^{k_{j+1}}\cdots\xi_n^{k_n}|dP.\end{aligned}$$

由所给条件知上述不等式右端每一项有限, 因而 $E|\xi_1^{k_1}\cdots\xi_n^{k_n}|<\infty$.

再者, (29) 与 (30) 的证明与性质 6 一样, 不再详述.

ii) 往证 (31) 式. 对不全为零的 $t_1,\cdots,t_n$, 令 $t=|t_1|+\cdots+|t_n|$, 选择 $u_1,\cdots,u_n$ 使 $\sum_{k=1}^{n}|u_k|=1$ 且 $u_kt=t_k,k=1,\cdots,n$. 于是对 $\xi=(\xi_1,\cdots,\xi_n)$,

$$f_\xi(t_1,\cdots,t_n)=f_\xi(tu_1,\cdots,tu_n)=f_{u_1\xi_1+\cdots+u_n\xi_n}(t).$$

由已知条件及 i) 知 $u_1\xi_1+\cdots+u_n\xi_n$ 的 m 阶绝对矩有限, 由性质 6 知

$$\begin{aligned}f_\xi(t_1,\cdots,t_n)&=f_{u_1\xi_1+\cdots+u_n\xi_n}(t)\\ &=\sum_{k=0}^{m-1}E(u_1\xi_1+\cdots+u_n\xi_n)^k\frac{(it)^k}{k!}+\rho_m(t).\end{aligned}$$

根据性质 6 中 (24) 式的第二尾式

$$\rho_m(t)=E(u_1\xi_1+\cdots+u_n\xi_n)^m\frac{(it)^m}{m!}+o(|t|^m)$$

$$= \sum_{\substack{k_1+\cdots+k_n=m\\ k_1,\cdots,k_n\geqslant 0}} \alpha_{k_1\cdots k_n}\frac{(it_1)^{k_1}}{k_1!}\cdots\frac{(it_n)^{k_n}}{k_n!}+o(|t|^m).$$

由 C_r 不等式知

$$|t|^2=||t_1|+\cdots+|t_n||^2\leqslant n(t_1^2+\cdots+t_n^2),$$

故知 (31) 具有 (32) 中第一尾式.

最后由 (24) 式第三尾式知

$$\begin{aligned}|\rho_m(t)| &\leqslant E|u_1\xi_1+\cdots+u_n\xi_n|^m\frac{|t|^m}{m!}\\ &\leqslant \sum_{\substack{k_1+\cdots+k_n=m\\ k_1,\cdots,k_n\geqslant 0}}\frac{|t_1|^{k_1}\cdots|t_n|^{k_n}}{k_1!\cdots k_n!}E|\xi_1^{k_1}\cdots\xi_n^{k_n}|,\end{aligned}$$

即 (31) 式具有 (32) 中第二尾式. 当 $t_1,\cdots,t_n$ 全为零时, 结论显然. □

(30) 式给出了计算多维随机变量的分布矩的方法, 下面以多维正态律为例.

例 14　设 $\xi=(\xi_1,\cdots,\xi_n)$ 是按 $N(m,D)$ 分布的, 则由例 8 知 ξ 的特征函数

$$f_\xi(t)=\exp\left\{it\cdot m'-\frac{1}{2}tDt'\right\}.$$

由例 10 知 ξ_k 按 $N(m_k,\sigma_k^2)$ 分布, 再由例 13 知 $E\xi_k=m_k$, $E\xi_k^2$ 有限, $k=1,2,\cdots n$. 再利用 Schuarz 不等式知 $E\xi_k\xi_l$ 有限.

再由 (30) 算得

$$E\xi_k\xi_l=i^{-2}\frac{\partial^2}{\partial t_k\partial t_l}f_\xi(t_1,\cdots,t_n)|_{t=0}=d_{kl}+m_km_l, k,l=1,\cdots,n.$$

因此

$$E(\xi_k-E\xi_k)(\xi_l-E\xi_l)=d_{kl},k,l=1,\cdots,n.$$

即 m 是 ξ 的数学期望, D 是 ξ 的相关方阵.

6.1.3　有限 L–S 测度的特征函数或 Fourier-Stieltjes 变式

为了更广泛地使用特征函数的方法, 我们把特征函数的定义加以推广.

定义 2　设 μ 是 $R^{(n)}$ 上任一有限 L–S 测度, 则称函数

$$f(t_1,\cdots,t_n)=\int e^{it\cdot x'}d\mu, t\in R^{(n)}$$

为测度 μ(或其分布函数) 的 Fourier-Stieltjes 变式或特征函数.

容易证明, 对任意有限 L–S 测度, 其特征函数是存在且有限的, 且有下述一系列性质, 其证明和 6.1.2 类似. 以下各性质中 $f(t)$ 均指 L–S 测度 μ 的特征函数.

性质 1′ a) $f(0)=\mu(R^{(n)})$;

b) $|f(t)|\leqslant f(0)$;

c) $\overline{f(t)}=f(-t)$.

性质 4′ 设 μ_k 是 $R^{(m_k)}$ 上的有限 L–S 测度, $f_k(t^{(m_k)})$ 是其特征函数, $k=1,\cdots,n.\mu=\mu_1\times\cdots\times\mu_n$ 是 $\mu_1,\cdots,\mu_n$ 的乘积测度, 则 μ 的特征函数

$$f(t)=\prod_{k=1}^{n}f_k(t^{(m_k)}),\text{其中}t=(t^{(m_1)},\cdots,t^{(m_n)})\in R^{(m_1+\cdots+m_n)}.$$

性质 5′(增量不等式).

$$|f(t)-f(t+h)|^2\leqslant 2f(0)[f(0)-Rf(h)]=2\mu(R^{(n)})[\mu(R^{(n)})-Rf(h)]$$

且 $f(t)$ 在 $R^{(n)}$ 上一致连续.

性质 6′ 在性质 6 中作如下变更: 设 μ 是 $R^{(1)}$ 上有限 L–S 测度. 令

$$\beta_r=\int|x|^r d\mu,\alpha_k=\int x^k d\mu.$$

则一切结论保持.

性质 7′ 设 μ 是 $R^{(n)}$ 上有限 L–S 测度. 在性质 7 中以 $\int|x_1^{r_1}\cdots x_n^{r_n}|d\mu$ (记作 $\beta_{r_1\cdots r_n}$) 代替 $E|\xi_1^{r_1}\cdots\xi_n^{r_n}|$, 以 $\int x_1^{k_1}\cdots x_n^{k_n}d\mu$ (记作 $\alpha_{k_1\cdots k_n}$) 代替 $E\xi_1^{k_1}\cdots\xi_n^{k_n}$, 则一切结论保持.

习题及补充

1. 试证: 超几何分布

$$P(\xi=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},k=0,1,\cdots$$

(其中 M,N 是实数, $N>M$) 的特征函数

$$f(t)=\frac{(N-M)(N-M-1)\cdots(N-M-n+1)}{N(N-1)\cdots(N-n+1)}\times F(-n,-M,N-M-n+1,e^{it}),$$

其中

$$F(\alpha,\beta,\gamma,t)=1+\frac{\alpha\beta}{\gamma}t+\frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)}\frac{t^2}{2!}+\cdots$$

是超几何函数 (这就是超几何分布命名的由来).

2. 试求下列分布的特征函数:

i) Γ- 分布: $p(x)=\begin{cases}\dfrac{c^r}{\Gamma(r)}x^{r-1}e^{-cx}, x>0,(r>0,c>0)\\ 0, \qquad\qquad x\leqslant 0;\end{cases}$

ii) Cauchy 分布: $p(x)=\dfrac{1}{\pi}\dfrac{a}{a^2+x^2}, a>0;$

iii) Laplace 分布: $p(x)=\dfrac{1}{2a}e^{-|x-b|/a}, a>0;$

iv) 负指数分布: $p(x)=\begin{cases}\alpha e^{-\alpha x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}\quad(\alpha>0).$

3. 试由特征函数的定义, 用视察法找出随机变量的分布律使它的特征函数就是相应的函数:

i) e^{iat};

ii) $\cos t$;

iii) $\cos^2 t$;

iv) $\sum\limits_{k=0}^{\infty}a_k e^{i\lambda_k t}, a_k\geqslant 0, \sum\limits_{k=0}^{\infty}a_k=1;$

v) $\sum\limits_{k=0}^{\infty}a_k\cos kt, a_k\geqslant 0, \sum\limits_{k=0}^{\infty}a_k=1;$

vi) $\dfrac{1}{1+it}$;(提示: 应用第 2 题 iv) 及性质 1.)

vii) $\dfrac{\sin at}{at}$.

4. 设 $\varphi(t), f(t)$ 是概率分布函数的特征函数, 且 $\varphi(t)$ 是实的. 试证:

i) $1-\varphi(2t)\leqslant 4(1-\varphi(t))$;

ii) $1-|f(2t)|^2\leqslant 4(1-|f(t)|^2)$. (提示: 首先证明 $|f(t)|^2$ 也是特征函数.)

5. 设 ξ 是离散型随机变量, $\sum\limits_{k=-\infty}^{\infty}P(\xi=a+kh)=1$, 则称 ξ 为按格子点分布的. 试证: ξ 按格子点分布的充分与必要条件是有一非零实数 t_0 存在, 使得 $|f_\xi(t_0)|=1$.

6. 设 $f(t)$ 是随机变量的特征函数, 并有一趋于零的非零序列 $t_1,t_2\cdots$, 使

$$|f(t_k)|=1, k=1,2,\cdots.$$

试证: 有一实数 α 存在使得

$$f(t)=e^{i\alpha t}.$$

称此分布律为退化律 (或单点分布).

7. 设 ξ 的 n 阶绝对矩有限, 试证

$$E(\xi - E\xi)^n = i^{-n}\frac{d^n}{dt^n}[e^{-it(E\xi)}f_\xi(t)]_{t=0}.$$

8. 设 $F_1(x), F_2(x)$ 是 $R^{(n)}$ 上有界分布函数, $f_1(t), f_2(t)$ 是它们相应的特征函数,

$$F(x) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}F_1(x-y)dF_2(y), x \in R^{(n)}.$$

试证:$F(x)$ 的特征函数 $f(t) = f_1(t)f_2(t), t \in R^{(n)}$.

9. 如果 $\{f_n\}$ 是 $R^{(n)}$ 上 L–S 测度序列的特征函数序列且 $f_n(t) \to g(t), t \in R^{(n)}, g$ 在零点处连续, 则 g 在 $R^{(n)}$ 上一致连续. (提示: 应用增量不等式.)

§6.2 逆转公式及唯一性定理

上节我们对 $R^{(n)}$ 上有限 L–S 测度定义了特征函数. 因而特征函数是由 L–S 测度唯一确定的. 现在我们要证明有限 L–S 测度被其特征函数唯一决定. 并且满足一定条件的分布函数也由其特征函数唯一确定. 这些统称为唯一性定理. 我们还将叙述唯一性定理在概率论中的某些应用.

6.2.1 逆转公式及唯一性定理

定理 1(逆转公式) 设 $f(t)$ 是有限 L–S 测度 μ 的特征函数, 若 $a, b \in R^{(n)}$ 且 $\mu([a,b] - (a,b)) = 0$, 则

$$\mu([a,b)) = \lim_{T\to\infty}\frac{1}{(2\pi)^n}\int_{-T}^{T}\cdots\int_{-T}^{T}\prod_{k=1}^{n}\frac{e^{-it_ka_k} - e^{-it_kb_k}}{it_k} \times f(t_1,\cdots,t_n)dt_1\cdots dt_n \tag{1}$$

等式右端的积分是 R- 积分.

证. i) 先证 (1) 式右端的积分对任何 $T > 0$ 都是存在的, 并且其被积函数连续有界, 因而可以认为它是 L- 积分.

因为对任何实数 c, d, t 来说, 当 $t \to 0$ 时,

$$\frac{e^{-itc} - e^{-itd}}{it} = e^{-i\frac{c+d}{2}t}2\frac{\sin(\frac{d-c}{2})t}{t} \to d - c,$$

$$\left|\frac{e^{-itc} - e^{-itd}}{it}\right| \leqslant |d - c|.$$

故 (1) 式右端被积函数连续有界, 因此可以将它看成 L 积分, 并且对任何 T 都存

在, 以下用 $I(T)$ 表示此积分. 而

$$I(T)=\int_{-T}^{T}\cdots\int_{-T}^{T}\left(\prod_{k=1}^{n}\frac{e^{-it_ka_k}-e^{-it_kb_k}}{it_k}\right)\times\left[\int_{R^{(n)}}e^{i\sum\limits_{k=1}^{n}t_kx_k}d\mu(x)\right]dt_1\cdots dt_n.$$

由于

$$\prod_{k=1}^{n}\frac{e^{-it_ka_k}-e^{-it_kb_k}}{it_k}e^{i\sum\limits_{k=1}^{n}t_kx_k}$$

是有界可测函数, 因而它在测度空间 $(R^{(n)}\times[-T,T],\mathscr{B}^{(n)}\times([-T,T]\cap\mathscr{B}^{(n)}),\mu\times\Lambda)$ 上是可积的, 其中 Λ 是 L 测度, 由 Fubini 定理, 这个积分就是 $I(T)$ 且

$$\begin{aligned}I(T)&=\int_{R^{(n)}}\int_{-T}^{T}\cdots\int_{-T}^{T}\prod_{k=1}^{n}\frac{e^{-it_ka_k}-e^{-it_kb_k}}{it_k}e^{i\sum\limits_{k=1}^{n}t_kx_k}\times dt_1\cdots dt_nd\mu(x)\\&=\int_{R^{(n)}}\prod_{k=1}^{n}\left[\int_{-T}^{T}\frac{e^{-it_ka_k}-e^{-it_kb_k}}{it_k}e^{it_kx_k}dt_k\right]d\mu(x).\end{aligned}\tag{2}$$

里层积分是 L 积分, 但由于被积函数连续, 可以看成 R 积分.

$$\begin{aligned}\prod_{k=1}^{n}\int_{-T}^{T}\frac{e^{-it_ka_k}-e^{-it_kb_k}}{it_k}e^{it_kx_k}dt_k&=2^n\prod_{k=1}^{n}\int_{0}^{T}\frac{\sin t_k(x_k-a_k)-\sin t_k(x_k-b_k)}{t_k}dt_k\\&=2^n\prod_{k=1}^{n}\int_{T(x_k-b_k)}^{T(x_k-a_k)}\frac{\sin t}{t}dt\end{aligned}$$

$$\xrightarrow{(T\to\infty)}\begin{cases}(2\pi)^n, & (x_1,\cdots,x_n)\in(a,b),\\2^{n-l}\pi^n, & (x_1,\cdots,x_n)\in[a,b]\backslash(a,b)\text{且恰有}l\text{个}k\text{使}x_k=a_k\text{或}b_k,\\0, & (x_1,\cdots,x_n)\notin[a,b].\end{cases}$$

上式最后的极限过程用到了下列的 Dirichlet 公式:

$$\lim_{x\to\infty}\int_0^x\frac{\sin t}{t}dt=\frac{\pi}{2}.$$

由于 $\int_0^x\frac{\sin t}{t}dt$ 在 $x\in R^{(1)}$ 上有界, 故 (2) 的里层积分对 T 和 $x_1,\cdots,x_n$ 一致有界, 因而由控制收敛定理及 $\mu([a,b]-(a,b))=0$ 知

$$\begin{aligned}\lim_{T\to\infty}I(T)&=\int_{(a,b)}(2\pi)^nd\mu+\int_{[a,b]^c}0d\mu\\&=(2\pi)^n\mu((a,b))=(2\pi)^n\mu([a,b)).\end{aligned}\tag{3}$$

这里用到了 $\mu([a,b]-(a,b))=0$. 于是 (1) 获证. □

定理 1 说明, 若 a,b 满足条件 $\mu([a,b]-(a,b))=0$, 则 $\mu([a,b))$ 的值由特征函数唯一决定. 为要证明对一切 $a,b\in R^{(n)}, a<b, \mu([a,b))$ 皆由其特征函数决定, 我们引入下述

定义 1 设 μ 是 $R^{(n)}$ 上 L–S 测度, 若 $a,b\in R^{(n)}$, 具有性质

$$\mu((a,b))=\mu([a,b]),$$

则称 $[a,b)$ 为μ **的一个连续区间**.

显然定理 1 中的 $[a,b)$ 是 μ 的连续区间. 于是为要证明有限 L–S 测度被其特征函数唯一决定, 只需证明任何一个 L–S 测度被其连续区间上的值唯一决定. 为此先证明以下几个引理.

引理 1 设 μ 是 $R^{(n)}$ 上 L–S 测度, 令

$$D(\mu)=\{a:a\in R^{(1)},\text{存在}k,1\leqslant k\leqslant n,\text{使}\mu(\{(x_1,\cdots,x_n):x_k=a\})>0\}. \tag{4}$$

则 $D(\mu)$ 是一可数集.

注 集合 $\{(x_1,\cdots,x_n):x_k=a\}$ 是与坐标轴平行的 $n-1$ 维超平面. 此引理说明: 与坐标轴平行的 $n-1$ 维超平面中至多有可数个测度大于零.

证 我们只需证: 对每一 $k,1\leqslant k\leqslant n$,

$$D_k=\{a:a\in R^{(1)},\mu(\{(x_1,\cdots,x_n):x_k=a\})>0\}$$

是可数集. 由于对不同的 a_1,a_2

$$\{(x_1,\cdots,x_n):x_k=a_1\}\cap\{(x_1,\cdots,x_n):x_k=a_2\}=\varnothing.$$

因此由第 1 章 §1.4 习题 3 知对任一 σ-有限测度来说, 测度大于零的不相交的集至多可数, 因而 D_k 是可数集 (包括有限和可数无限两种情形).

引理 2 设 μ 是 $R^{(n)}$ 上的 L–S 测度, $D(\mu)\subset R^{(1)}$ 是由 (4) 定义的可数集, 令

$$C(\mu)=R^{(1)}-D(\mu). \tag{5}$$

若 $a,b\in R^{(n)}$, 且 a,b 的每一坐标 $a_k,b_k\in C(\mu)$, 则$[a,b)$ 是 μ 的连续区间.

证 由于

$$[a,b]-(a,b)\subset\left(\bigcup_{k=1}^{n}\{(x_1,\cdots,x_n):x_k=a_k\}\right)\bigcup\left(\bigcup_{k=1}^{n}\{(x_1,\cdots,x_n):x_k=b_k\}\right),$$

而 $a_k, b_k \notin D(\mu), k=1,\cdots,n$. 因而上式右端每一个集合的 μ 测度为零. 所以

$$\mu([a,b]-(a,b))=0.$$

即 $[a,b)$ 为 μ 的一个连续区间. □

引理 3　设 μ_1,μ_2 是 $R^{(n)}$ 上两个 L–S 测度, 且对 μ_1 和 μ_2 的任何共同的连续区间来说, μ_1 与 μ_2 在它上面的值相等, 则 $\mu_1=\mu_2$.

这一引理说明任一 L–S 测度由它在连续区间上的值唯一决定.

证　由于 $D(\mu_1)\cup D(\mu_2)$ 是 $R^{(1)}$ 中的可数集, 因而

$$C(\mu_1)\cap C(\mu_2)=R^{(1)}-(D(\mu_1)\cup D(\mu_2))$$

在 $R^{(1)}$ 中稠. 于是对任何 $a,b\in R^{(n)}, a<b$ 及充分小的 $\varepsilon>0$, 总存在 $[a_\varepsilon,b_\varepsilon)$, 其 2^n 个顶点的各个坐标均在 $C(\mu_1)\cap C(\mu_2)$ 中, 且满足

$$a<a_\varepsilon<a+\varepsilon<b-\varepsilon<b_\varepsilon<b.$$

因而 $[a_\varepsilon,b_\varepsilon)$ 是 μ_1,μ_2 的共同连续区间. 由此即知可取一区间序列

$$[a^{(k)},b^{(k)}), k=1,2,\cdots$$

使得它满足以下条件:

a) 每一 $[a^{(k)},b^{(k)})$ 是 μ_1,μ_2 的共同连续区间;

b) $[a^{(1)},b^{(1)})\subset[a^{(2)},b^{(2)})\subset\cdots$;

c) $\bigcup\limits_{k=1}^{\infty}[a^{(k)},b^{(k)})=(a,b)$.

于是由 μ_1,μ_2 的连续性知

$$\mu_1((a,b))=\lim_{k\to\infty}\mu_1([a^{(k)},b^{(k)}))=\lim_{k\to\infty}\mu_2([a^{(k)},b^{(k)}))=\mu_2((a,b)).$$

由于 a,b 的任意性, 对任何 $a,b\in R^{(n)}$,

$$\begin{aligned}\mu_1([a,b))&=\lim_{m\to\infty}\mu_1\left(\left(a-\frac{1}{m},b\right)\right)\\&=\lim_{m\to\infty}\mu_2\left(\left(a-\frac{1}{m},b\right)\right)=\mu_2([a,b)).\end{aligned}$$

而 L–S 测度由其有限区间上的值唯一确定, 因而

$$\mu_1(B)=\mu_2(B), B\in\mathscr{B}^{(n)}.$$ □

由定理 1 和引理 3 我们就得到

定理 2 有限 L–S 测度被其特征函数唯一决定.

为了讨论什么样的分布函数被其特征函数唯一决定, 我们引入

定义 2 若 F 是 $R^{(n)}$ 上有界不降分布函数, 且对一切 $1 \leqslant k \leqslant n, 1 \leqslant i_1 < \cdots < i_k \leqslant n$及一切$x_{j_1}, \cdots, x_{j_{n-k}} \in R^{(1)}((i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$是$(1, \cdots, n)$ 的一个置换), 有

$$\lim_{\substack{x_{i_s} \to -\infty \\ s=1,\cdots,k}} F(x_1, \cdots, x_n) = 0, \tag{6}$$

则称 F 为标准分布函数.

性质 1 若 μ 是有限 L–S 测度, 则

$$F(x) = \mu((-\infty, x)), x \in R^{(n)}$$

是标准分布函数.

性质 2 若 F 是标准分布函数, μ_F 是与其对应的 L–S 测度, 则

$$F(x) = \mu_F((-\infty, x)), x \in R^{(n)}. \tag{7}$$

证 对一切 $x \in R^{(n)}, a < x$

$$\Delta_{x,a} F = \mu_F([a, x)).$$

取递降序列 $a^{(k)} \downarrow -\infty$, 由$F$ 的标准性及 μ 的连续性知

$$F(x) = \lim_{k \to \infty} \Delta_{x,a^{(k)}} F = \lim_{k \to \infty} \mu_F([a^{(k)}, x)) = \mu_F((-\infty, x)). \qquad \square$$

定理 3 标准分布函数由其特征函数唯一决定. 特别, 概率分布函数由其特征函数唯一决定.

证 设 F 是标准分布函数, μ_F 是与其对应的 L–S 测度, 则 μ_F 有限, 且 F 的特征函数也是 μ_F 的特征函数. 由定理 2 知 μ_F 由其特征函数唯一决定, 而 F 又由 μ_F 唯一决定, 因而标准分布函数由其特征函数唯一决定.

因为概率分布函数是标准分布函数, 因此后一结论是前一结论的特殊情形. □

关于有限 L–S 测度 (或标准分布函数) 由其矩唯一决定的问题见习题 7.

6.2.2 唯一性定理在概率论中的应用

逆转公式及唯一性定理在概率论中有广泛的应用, 今举出几点加以说明.

1° 应用上述定理可以得出一系列各种重要的具体分布律的性质:

I. 关于正态律的性质. 由例 8 知 $N(m, D)$ 的特征函数

$$f(t) = \exp\{e^{it \cdot m' - \frac{1}{2} tDt'}\}. \tag{8}$$

又由定理 3 知概率分布函数由其特征函数唯一决定, 故有

推论 1　若 n 维随机变量 ξ 的特征函数具有 (8) 的形状, 则 ξ 是按 $N(m,D)$ 分布的.

其次, 由推论 1 及 §6.1 例 9 知

推论 2　若 $\xi=(\xi_1,\cdots,\xi_n)$ 按 $N(m,D)$ 分布, 则 $\xi_1\cdots+\xi_n$ 按 $N\left(\sum\limits_{k=1}^{n}m_k, \sum\limits_{k,l=1}^{n}d_{kl}\right)$ 分布, 特别若 ξ_k 按 $N(m_k,\sigma_k^2)$ 分布, $k=1,\cdots,n$, 且独立, 则 $\sum\limits_{k=1}^{n}\xi_k$ 按 $N\left(\sum\limits_{k=1}^{n}m_k,\sum\limits_{k=1}^{n}\sigma_k^2\right)$ 分布. 一般地, 若 $\xi^{(k)}=(\xi_1^{(k)},\cdots,\xi_r^{(k)}),k=1,\cdots,n$ 独立, 且 $\xi^{(k)}$ 按 $N(m^{(k)},D^{(k)})$ 分布, 则 $\sum\limits_{k=1}^{n}\xi^{(k)}$ 按 $N\left(\sum\limits_{k=1}^{n}m^{(k)},\sum\limits_{k=1}^{n}D^{(k)}\right)$ 分布.

第三, 由 §6.1 例 10 知

推论 3　若 $\xi=(\xi_1,\cdots,\xi_n)$ 按 $N(m,D)$ 分布, 则 $(\xi_{i_1},\cdots,\xi_{i_k})(1\leqslant k\leqslant n,1\leqslant i_1<\cdots<i_k\leqslant n)$ 按 $N(m^{(1)},D^{(1)})$ 分布, 其中 $m^{(1)}=(m_{i_1},\cdots,m_{i_k}),D^{(1)}=(d_{i_s i_l})_{1\leqslant s,l\leqslant k}$.

以前我们总假定 $N(m,D)$ 分布中的 $n\times n$ 方阵 D 是正定的, 此时它的特征函数具有 (8) 的形状. 但是由 §6.4 将要讨论的特征函数极限定理知, 若 D 是非负定方阵 (即 D 所对应的二次型是非负的), 则形如 (8) 的函数仍为某一 n 维随机变量的特征函数. 因此我们将正态律的概念推广成

定义 3　若 D 为一 $n\times n$ 非负定方阵, 则称具有形如 (8) 的特征函数的随机变量 ξ 为按 n 维正态律分布的, 记作 $\xi\sim N(a,D)$. 若 $|D|\neq 0$, 称 $N(a,D)$ 为**满秩的 n 维正态律**; 若 $|D|=0$, 则称为**降秩的 n 维正态律.**

因此, 我们以前讨论的 n 维正态律是满秩的正态律.

由 §6.1 性质 3 易知

推论 4　设 $A=(a_{lk})$ 为 $m\times n$ 矩阵, $b=(b_1,\cdots,b_m)$. 若 $\xi=(\xi_1,\cdots,\xi_n)$ 按 $N(a,D)$ 分布, 则 $\eta=\xi A'+b$ 按 n 维正态律 $N(aA'+b,ADA')$ 分布.

若 $m\leqslant n$, 且 A 的秩为 $m,|D|\neq 0$, 则 η 按满秩的 m 维正态律分布.

下面举一个应用的例子.

例 1　设 $\xi_1,\cdots,\xi_n$ 是 n 个按 $N(0,1)$ 分布的独立随机变量,

$$\bar{\xi}=\frac{1}{n}\sum_{k=1}^{n}\xi_k,\quad s^2=\frac{1}{n-1}\sum_{k=1}^{n}(\xi_k-\bar{\xi})^2.$$

则 $\bar{\xi}$ 与 s^2 是独立随机变量且 $(n-1)s^2$ 按自由度为 $n-1$ 的 χ^2-律分布. $\bar{\xi}$ 按 $N\left(0,\dfrac{1}{n}\right)$

分布, 因而

$$t = \frac{\sqrt{n}\bar{\xi}}{s}$$

按自由度为 $n-1$ 的 t- 分布律分布.

证 容易算出 $\sum_{k=1}^{n}(\xi_k - \bar{\xi})^2 = \sum_{k=1}^{n}\xi_k^2 - n\bar{\xi}^2$, 由于 $(\xi_1, \cdots, \xi_n)$ 按 $N(0, I^{(n)})$ 分布, 其中 $I^{(n)}$ 为 n 阶单位方阵,

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} 0 \cdots 0 \\ \frac{1}{\sqrt{2\times 3}} \frac{1}{\sqrt{2\times 3}} \frac{-2}{\sqrt{2\times 3}} \cdots 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{1}{\sqrt{n(n-1)}} \frac{1}{\sqrt{n(n-1)}} \cdots \frac{1}{\sqrt{n(n-1)}} \frac{-(n-1)}{\sqrt{n(n-1)}} \\ \frac{1}{\sqrt{n}} \frac{1}{\sqrt{n}} \cdots \frac{1}{\sqrt{n}} \frac{1}{\sqrt{n}} \end{pmatrix}$$

是一 $n \times n$ 正交方阵, 于是按推论 4, $\eta = \xi A'$ 按 $N(0, I^{(n)})$ 分布, 因此 $\eta_1, \cdots, \eta_n$ 是按 $N(0,1)$ 分布的独立随机变量, 且由 $AA' = I^{(n)}$ 及 $\eta = \xi A'$ 知

$$\eta_n = \frac{1}{\sqrt{n}}(\xi_1 + \cdots + \xi_n) = \sqrt{n}\bar{\xi};$$

$$\sum_{k=1}^{n}\xi_k^2 = \xi \cdot \xi' = (\eta A)(A'\eta') = \eta \cdot \eta' = \sum_{k=1}^{n}\eta_k^2.$$

因而由第 3 章 §3.4 知

$$\sum_{k=1}^{n}(\xi_k - \bar{\xi})^2 = \sum_{k=1}^{n}\xi_k^2 - n\bar{\xi}^2 = \sum_{k=1}^{n}\eta_k^2 - \eta_n^2 = \sum_{k=1}^{n-1}\eta_k^2$$

与 η_n 独立, 于是 s^2 与 $\bar{\xi}$ 独立. 再由 χ^2-分布律的定义知 $\sum_{k=1}^{n}(\xi_k - \xi)^2$ 按自由度为 $n-1$ 的 χ^2-律分布. 最后按 t- 分布的定义知 $\frac{\sqrt{n}\bar{\xi}}{s}$ 按 t- 分布律分布.

由此容易推出, 若 $\xi_1, \cdots, \xi_n$ 是按 $N(a, \sigma^2)$ 分布的独立随机变量, 则 $\bar{\xi}$ 与 $\frac{1}{\sigma^2}\sum_{k=1}^{n}(\xi_k - \bar{\xi})^2$ 独立, 且后者按自由度为 $n-1$ 的 χ^2-律分布, 且

$$\frac{\sqrt{n}(\bar{\xi} - a)}{s}$$

按自由度为 $n-1$ 的 t- 分布律分布.

II. 关于 χ^2-律的性质. 由 §6.1 例 6 知自由度为 n 的 χ^2-律的特征函数为

$$(1-2it)^{-\frac{n}{2}}. \tag{9}$$

于是由定理 3 知具有形如 (9) 的特征函数的标准分布函数必为自由度为 n 的 χ^2-律的分布函数. 因此由 §6.1 性质 4 得

推论 5　设 $\chi_1^2, \cdots, \chi_r^2$ 分别是按自由度为 $n_k, k=1,\cdots,r$ 的 χ^2-律分布的独立随机变量, 则 $\sum\limits_{k=1}^{r}\chi_k^2$ 为按自由度为 $\sum\limits_{k=1}^{r}n_k$ 的 χ^2-律分布的随机变量.

由 §6.1 性质 4、例 6 及定理 3 还立刻得到下面的

推论 6　设 χ^2, χ_1^2 分别为按自由度为 n, n_1 的 χ^2-律分布的随机变量, $n_1 < n, \chi_1^2+\xi=\chi^2, \chi_1^2$ 与 ξ 独立, 则 ξ 是按自由度为 $n-n_1$ 的 χ^2-律分布的随机变量.

证　由 §6.1 例 6 及性质 4 知

$$(1-2it)^{-\frac{n}{2}} = (1-2it)^{-\frac{n_1}{2}} f_\xi(t).$$

于是

$$f_\xi(t) = (1-2it)^{-\frac{n-n_1}{2}}.$$

因此, 由 §6.1 例 6 及定理 3 即得所需结论.

2°　应用定理 3 可以得到实随机变量独立的判别条件.

定理 4　设 $\xi^{(k)} = (\xi_1^{(k)}, \cdots, \xi_{m_k}^{(k)}), k=1,\cdots,n$ 是 n 个随机变量, $\xi = (\xi_1^{(1)}, \cdots, \xi_{m_1}^{(1)}, \cdots, \xi_1^{(n)}, \cdots, \xi_{m_n}^{(n)})$, 若 $\xi^{(k)}, k=1,\cdots,n$ 独立, 则 ξ 的特征函数

$$f_\xi(t) = f_\xi^{(1)}(t^{(1)}) \cdots f_\xi^{(n)}(t^{(n)}). \tag{10}$$

$$t = (t^{(1)}, \cdots, t^{(n)}), t^{(k)} = (t_1^{(k)}, \cdots, t_{m_k}^{(k)}) \in R^{(m_k)}.$$

反之, 若 ξ 的特征函数

$$f_\xi(t) = f_1(t^{(1)}) \cdots f_n(t^{(n)}), t = (t^{(1)}, \cdots, t^{(n)}), \tag{11}$$

$$t^{(k)} = (t_1^{(k)}, \cdots, t_{m_k}^{(k)}) \in R^{(m_k)},$$

则 $\xi^{(k)}, k=1,\cdots,n$, 独立且 $\xi^{(k)}$ 的特征函数

$$f_\xi^{(k)}(t^{(k)}) = f_k(t^{(k)})/f_k(0^{(k)}), t^{(k)} \in R^{(m_k)}, \tag{12}$$

其中 $0^{(k)} = (0, \cdots, 0) \in R^{(m_k)}$.

证 定理的第一部分即 §6.1 性质 4 的第一部分. 今证定理的第二部分. 由于 $m_k > 1$ 的情形的证明与 $m_k = 1$ 的情形的证明完全一样, 因而只写出 $m_k = 1$ 的情形的证明.

首先, 由 §6.1 性质 2 及 (11) 知

$$f_{\xi_k}(t_k) = f_\xi(0,\cdots,0,t_k,0,\cdots,0) = f_k(t_k)\prod_{j\neq k} f_j(0),$$

再由

$$f_\xi(0,\cdots,0) = \prod_{k=1}^{n} f_k(0) = 1$$

知

$$\prod_{j\neq k} f_j(0) = \frac{1}{f_k(0)}.$$

于是 (12) 式获证.

为了简单, 以下不妨设 $f_k(0) = 1$, 即 $f_k(t)$ 是 ξ_k 的特征函数.

若 $[a,b)$ 的 2^n 个顶点的所有坐标皆属于 $C(P_\xi)$, 则它显然是 P_ξ 的连续区间, 而且由 P_ξ 在这一类区间上的值完全决定了 P_ξ. 而且 $[a_k,b_k)$ 也是 P_{ξ_k} 的连续区间, $k=1,\cdots,n$. 由定理 1 及 Fubini 定理

$$\begin{aligned}\Delta_{b,a}P_\xi = P_\xi([a\quad b)) &= \lim_{T\to\infty}\frac{1}{(2\pi)^n}\int_{-T}^{T}\cdots\int_{-T}^{T}\\ &\quad \prod_{k=1}^{n}\frac{e^{-it_ka_k}-e^{-it_kb_k}}{it_k}f_k(t_k)dt_1\cdots dt_n\\ &= \prod_{k=1}^{n}\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-it_ka_k}-e^{-it_kb_k}}{it_k}f_k(t_k)dt_k\\ &= \prod_{k=1}^{n}\Delta_{b_k,a_k}F_{\xi_k}. \end{aligned}\tag{13}$$

令

$$F'(x) = \prod_{k=1}^{n} F_{\xi_k}(x),\tag{14}$$

显然 $F'(x)$ 是概率分布函数且

$$\Delta_{b,a}F' = \prod_{k=1}^{n}\Delta_{b_k,a_k}F_{\xi_k},$$

即对一切 $[a,b)$ 的 2^n 个顶点的每一坐标均属于 $C(P_\xi)$ 的 $[a,b)$,

$$\Delta_{b,a}F_\xi = \Delta_{b,a}F'.\tag{15}$$

仿引理 3 的证明由 (15) 知

$$F_\xi(x) = F'(x) = \prod_{k=1}^{n} F_{\xi_k}(x), x \in R^{(n)}.$$

再由第 2 章 §2.4 知 $\xi_1, \cdots, \xi_n$ 独立. □

(11) 式也是一个使用方便的独立判别条件.

习题及补充

1. 求证：函数

$$f_1(t) = \frac{1}{\mathrm{ch}t}; f_2(t) = \frac{t}{\mathrm{sh}t}; f_3(t) = \frac{1}{\mathrm{ch}^2 t}$$

各为具有下列分布密度的随机变量的特征函数：

$$p_1(x) = \frac{1}{2\mathrm{ch}\frac{\pi x}{2}}; p_2(x) = \frac{\pi}{4\mathrm{ch}^2\frac{\pi x}{2}}; p_3(x) = \frac{x}{2\mathrm{sh}\frac{\pi x}{2}}.$$

2. 若 F_1, F_2, 为标准分布函数, f_1, f_2 分别是它们所对应的特征函数, $f(t) = f_1(t)f_2(t)$, 则 $f(t)$ 对应的标准分布函数

$$F(x) = \int_{-\infty}^{\infty} F_1(x-y)dF_2(y).$$

3. 设 $F(x)$ 是 $R^{(1)}$ 上有界分布函数, $f(t)$ 是 $F(x)$ 的特征函数, 试证：对 $F(x)$ 的任意连续点 $a, b, (a < b)$,

$$F(b) - F(a) = \frac{1}{\pi}\int_0^{\infty} \frac{I\{(e^{-ita} - e^{-itb})f(t)\}}{t} dt,$$

其中 Iz 表示复数 z 的虚部.

4. 设 $R^{(1)}$ 上有界分布函数 $F(x)$ 的特征函数是 $f(t)$, 试证：

i) 对任何实数 x 恒有

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} f(t)e^{-itx}dt = F(x+0) - F(x-0);$$

ii) 若 $\{x_\nu\}$ 是 $F(x)$ 的一切间断点的序列, 则

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |f(t)|^2 dt = \sum_{\nu}\{F(x_\nu + 0) - F(x_\nu - 0)\}^2.$$

5. 设标准分布函数 $F(x)$ 的特征函数 $f(t)$ 对 Lebesgue 测度可积, 试证 $F(x)$ 的导数存在且

$$F'(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx} f(t)dt,$$

即 $F'(x)$ 是 $f(t)$ 的 Fourier 变换. F' 和 f 互为 Fourier 变换对.

6. 若标准分布函数是绝对连续的, 即

$$F(x) = \int_{-\infty}^{x} F'(t)dt,$$

则其特征函数 $f(t)$ 在 $|t| \to \infty$ 时趋于零.

7. 若标准分布函数 $F(x)$ 的各阶矩 $\{\alpha_n\}$ 都有限, 且 $\sum\limits_{n=0}^{\infty} \frac{\alpha_n}{n!} t^n$ 具有正的收敛半径 c, 则这个分布函数 $F(x)$ 由其各阶矩唯一决定. (提示: 只要证 $F(x)$ 的特征函数 $f(t)$ 由 $\{\alpha_k\}$ 唯一决定. 首先利用 6.1.2 性质 6 的推论知: 当 $u \in (-c, c)$ 时 $f(u) = \sum\limits_{k=0}^{\infty} \frac{\alpha_k}{k!}(iu)^k$. 其次, 利用

$$\begin{aligned} f(t+u) &= \int_{-\infty}^{\infty} e^{i(t+u)x} dF(x) \\ &= \int_{-\infty}^{\infty} \left[1 + \sum_{r=1}^{n-1} \frac{(iu)^r}{r!} x^r + \frac{(iu)^n}{n!} x^n (\cos u'x + i \sin u''x)\right] \\ &\quad \times e^{itx} dF(x) \end{aligned}$$

及证明 6.1.2 性质 6 推论的方法证明

$$f(t+u) = f(t) + \sum_{k=1}^{\infty} \frac{u^k}{k!} f^{(k)}(t), u \in (-c, c).$$

从而证明了 $f(t)$ 在 $t \in (-2c, 2c)$ 的值由 $\{\alpha_k\}$ 唯一决定. 使用归纳法可证对一切自然数 n, $f(t)$ 在 $t \in (-nc, nc)$ 的值由 $\{\alpha_k\}$ 唯一决定. 因而获证.)

8. 若 ξ 是有界随机变量, 试证其分布函数 $F(x)$ 由其一切阶矩 $\{\alpha_k\}$ 决定.

9. 设 ξ 为一实随机变量. 若对任一实数 x,

$$P(\xi < x) = P(\xi > -x),$$

则称 ξ 为对称随机变量. 试证 ξ 是对称随机变量的充分必要条件是 $f_\xi(t)$ 为实函数.

10. 如果特征函数满足 $f(t) = 1 + o(t^2), t \to 0$, 则$f(t) \equiv 1$.

§6.3 L–S 测度的弱收敛

为了研究特征函数的收敛性与分布律的某种收敛性之间的关系, 我们先来研究一致有界 L–S 测度的弱收敛. 关于距离空间中测度弱收敛的理论有许多学者进行研究, 例如在 [11] 中有较深入的结果. 我们在此只讨论一种特殊情形, 即 $R^{(n)}$ 上一致有界 L–S 测度的弱收敛性, 但方法是带有一般性的.

6.3.1 测度弱收敛的定义

定义 1　设 A 是 $R^{(n)}$ 的子集, $d(x,y)=\sqrt{(x-y)\cdot(x-y)'}$ 是 $R^{(n)}$ 中点 x,y 之间的距离,

i) 若 $x\in A$ 且存在 $\varepsilon>0$, 使

$$\{y:d(x,y)<\varepsilon\}\subset A,$$

则称x **为** A **的内点**, 称 A 的一切内点的集合为A **的内核**, 记作 A°;

ii) 若 $x\in R^{(n)}$ 且对一切 $\varepsilon>0$ 有

$$\{y:d(x,y)<\varepsilon\}\cap A\neq\varnothing, \tag{1}$$

则称 x 为 A**的极限点**, 称 A 的一切极限点的集合为 A**的闭包**, 记作 $\overline{A}$.

iii) 称集合 $\overline{A}-A^\circ$ 为 A 的边界, 记作 ∂A.

iv) 若 $A^\circ=A$, 称 A 为**开集**, 若 A^c 是开集称 A 为**闭集**.

定义 2　设 $\{\mu_k\}$ 是一致有界 L–S 测度序列, 若存在 L–S 测度 μ, 使对一切 μ 的连续区间 I 均有

$$\mu_k(I)\to\mu(I) \tag{2}$$

且

$$\mu_k(R^{(n)})\to\mu(R^{(n)}). \tag{3}$$

则称μ_k **弱收敛向** μ, 记作 $\mu_k\xrightarrow{\text{w}}\mu$.

由 §6.2 引理 3 知, 若 $\mu_k\xrightarrow{\text{w}}\mu,\mu_k\xrightarrow{\text{w}}\bar{\mu}$, 则 $\mu=\overline{\mu}$, 即弱收敛的极限唯一.

对一维情形, $R^{(1)}$ 上 L–S 测度序列 $\mu_k\xrightarrow{\text{w}}\mu$ 等价于在 $F(x)$ 的连续点上 $F_k(x)\to F(x)$且$F_k(+\infty)\to F(+\infty)$.其中$F_k,F$ 分别是 μ_k,μ 所对应的标准分布函数, 而 $F_k(+\infty)=\lim\limits_{x\to\infty}F_k(x),F(+\infty)=\lim\limits_{x\to\infty}F(x)$. 对概率测度而言, 分布律弱收敛即第 2 章 §2.5 所述的分布律收敛.

测度弱收敛有一系列等价形式. 为了叙述这些, 我们先证明

引理 1　设 μ_k 是 $R^{(n)}$ 上一致有界 L–S 测度序列, μ 是 L–S 测度, 若对一切 μ 连续区间 I 都有 $\mu_k(I)\to\mu(I)$, 则对任何 $R^{(n)}$ 上的有界连续函数 g 及任何 μ 连续区间 I 均有

$$\int_I gd\mu_k\to\int_I gd\mu. \tag{4}$$

证　设 $|\mu_k|<M,k=1,2,\cdots,|g|<N.g$在$I$ 的闭包 $\overline{I}$ 上一致连续, 因而任给 $\varepsilon>0$, 总存在 $\delta>0$, 当 $d(x,x')<\delta$时$|g(x)-g(x')|<\varepsilon$.

再由 §6.2 引理 1*知存在有限个 μ 测度为零的 $n-1$ 维超平面, 把区间 I 划分成直径小于 δ 的小区间, 设为 $I_1, I_2, \cdots, I_m$, 这些区间都是 μ 连续区间. 设 $I_j = [a^{(j)}, b^{(j)})$. 令

$$g^*(x) = \sum_{j=1}^{m} g(a^{(j)})\chi_{I_j}(x), x \in R^{(n)},$$

则 g^* 是有界 Borel 可测函数, 于是

$$\begin{aligned}\left|\int_I g d\mu - \int_I g d\mu_k\right| &\leqslant \int_I |g-g^*|d\mu + \left|\int_I g^* d\mu - \int_I g^* d\mu_k\right| + \int_I |g-g^*|d\mu_k \\ &\leqslant \varepsilon M + \sum_{j=1}^{m} |g(a^{(j)})||\mu(I_j) - \mu_k(I_j)| + \varepsilon M \\ &\leqslant 2\varepsilon M + N\sum_{j=1}^{m} |\mu(I_j) - \mu_k(I_j)|.\end{aligned}$$

由于每一 I_j 是 μ 连续区间, 故当 $k \to \infty$ 时

$$\overline{\lim_{k\to\infty}}\left|\int_I g d\mu - \int_I g d\mu_k\right| \leqslant 2\varepsilon M.$$

再由 ε 的任意性, 知

$$\overline{\lim_{n\to\infty}}\left|\int_I g d\mu - \int_I g d\mu_k\right| = 0.$$

故引理获证. □

定理 1 设 $\mu_k \xrightarrow{\text{w}} \mu, g$ 是 $R^{(n)}$ 上有界连续函数, 则

$$\int g d\mu_k \to \int g d\mu. \tag{5}$$

证 设 $|g| < N$, 由于 μ 是有限测度, 并由 §6.2 引理 1 知, 存在 μ 连续区间 I 使

$$\mu(R^{(n)} - I) < \frac{\varepsilon}{6N},$$

再由 $\mu_k \xrightarrow{\text{w}} \mu$ 知 $\mu_k(R^{(n)}) \to \mu(R^{(n)})$ 及 $\mu_k(I) \to \mu(I)$. 故当 k 充分大时

$$\mu_k(R^{(n)} - I) < \frac{\varepsilon}{3N}.$$

由引理 1 知

$$\int_I g d\mu_k \to \int_I g d\mu,$$

故当 k 充分大时

$$\left|\int_I g d\mu - \int_I g d\mu_k\right| < \frac{\varepsilon}{3}.$$

于是对于充分大的 k,

$$\left|\int gd\mu-\int gd\mu_k\right|\leqslant\int_{I^c}|g|d\mu+\left|\int_I gd\mu-\int_I gd\mu_k\right|+\int_{I^c}|g|d\mu_k$$
$$\leqslant N\mu(I^c)+\varepsilon/3+N\mu_k(I^c)<\varepsilon. \tag{6}$$

定理获证 □

推论 设 $|\mu_k|<M, k=1,2,\cdots$, 在一切 μ 连续区间 I 上 $\mu_k(I)\to\mu(I), g$ 在 $R^{(n)}$ 上连续, 且当 $||x||=\sqrt{x\cdot x'}\to\infty$ 时 $g(x)$ 一致地趋于零, 则

$$\int gd\mu_k\to\int gd\mu.$$

证 由于 $||x||\to\infty$ 时 $g(x)$ 一致地趋于零, 故任给 $\varepsilon>0$, 存在 μ 连续区间 I, 使得当 $x\in I^c$ 时 $|g(x)|<\dfrac{\varepsilon}{3M}$, 因而和定理 1 中证明 (6) 式一样证得此推论. □

下面我们只就 L–S 测度来证明测度弱收敛的一系列等价形式. 一般距离空间上的测度弱收敛, 证明方法类似.

定理 2 设 $\{\mu_k\}$ 是一致有界 L–S 测度序列, μ 是 L–S 测度, 则下列命题等价:

i) $\mu_k\xrightarrow{\text{w}}\mu$;

ii) 对 $R^{(n)}$ 上一切有界连续函数 g(记作 $g\in C(R^{(n)})$) 有 $\int gd\mu_k\to\int gd\mu$;

iii) 对 $R^{(n)}$ 上一切有界一致连续函数 g(记作 $g\in U(R^{(n)})$), $\int gd\mu_k\to\int gd\mu$;

iv) $\overline{\lim\limits_{k\to\infty}}\mu_k(C)\leqslant\mu(C)$, 对一切闭集 C;

$$\underline{\lim_{k\to\infty}}\ \mu_k(G)\geqslant\mu(G), \text{对一切开集 } G;$$

v) $\lim\limits_{k\to\infty}\mu_k(A)=\mu(A)$, 若 $\mu(\partial A)=0, A\in\mathscr{B}^{(n)}$.

注 1 $R^{(n)}$ 中的开集是 Borel 集, 因而闭集以及 Borel 集的边界也是 Borel 集, 请读者自证.

注 2 若 μ_k,μ 均为概率测度, 则 iv) 中两个不等式等价, 均可作为 $\mu_k\xrightarrow{\text{w}}\mu$ 的充分必要条件.

证 i)⇒ii) 为定理 1 所证;

ii)⇒iii) 是由于 $C(R^{(n)})\supset U(R^{(n)})$;

往证 iii)⇒iv): 对 $R^{(n)}$ 中任一闭集 C, 考虑函数

$$d(x,C)\hat{=}\inf_{y\in C}d(x,y).$$

它是 $R^{(n)}$ 上一致连续函数 (请读者自证). 设

$$G_m=\left\{x:d(x,C)<\frac{1}{m}\right\}.$$

则 C 和 G_m^c 是不交闭集. 若令

$$f_m(x)=\frac{d(x,C)}{d(x,C)+d(x,G_m^c)},$$

则

$$0\leqslant|f_m(x)|\leqslant 1,x\in R^{(n)}$$

且当 $x\in G_m^c$ 时 $f_m(x)=1$, 当 $x\in C$ 时 $f_m(x)=0$. 由于

$$\inf_{\substack{x\in G_m^c\\ y\in C}} d(x,y)\geqslant\frac{1}{m},$$

易证对一切 m,f_m 是一致连续的, 且

$$\chi_C\leqslant 1-f_m\leqslant\chi_{G_m}.$$

再者, 易证 $G_1\supset G_2\supset\cdots$ 且 $\bigcap\limits_m G_m=C$. 于是对一切正整数 m 有

$$\overline{\lim_{k\to\infty}}\mu_k(C)\leqslant\lim_{k\to\infty}\int(1-f_m)d_{\mu_k}.$$

由 iii) 及 $1-f_m\in U(R^{(n)})$ 知

$$\lim_{k\to\infty}\int(1-f_m)d\mu_k=\int(1-f_m)d\mu\leqslant\mu(G_m),$$

即

$$\overline{\lim_{k\to\infty}}\mu_k(C)\leqslant\mu(G_m).$$

令 $m\to\infty$, 由测度的连续性知

$$\overline{\lim_{k\to\infty}}\mu_k(C)\leqslant\mu\left(\bigcap_m G_m\right)=\mu(C).$$

故 iv) 中第一式获证.

往证 iv) 中第二式: 对任何开集 G,G^c 是闭集, 由第一式知

$$\overline{\lim_{k\to\infty}}\mu_k(G^c)\leqslant\mu(G^c),$$

即

$$\overline{\lim_{k\to\infty}}[\mu_k(R^{(n)})-\mu_k(G)]\leqslant\mu(R^{(n)})-\mu(G).$$

由 iii) 取 $g=1$ 有

$$\lim_{k\to\infty}\mu_k(R^{(n)})=\lim_{k\to\infty}\int 1d\mu_k=\int 1d\mu=\mu(R^{(n)}).$$

故有

$$\underline{\lim_{k\to\infty}}\ \mu_k(G)\geqslant\mu(G).$$

故 iii)⇒iv) 获证.

往证 iv)⇒v)：设 $A\in\mathscr{B}^{(n)}$且$\mu(\partial A)=0$, 因为

$$A^\circ\subset A\subset\overline{A}$$

且

$$\mu(A^\circ)=\mu(A)=\mu(\overline{A}),$$

于是

$$\begin{aligned}\mu(A)=\mu(A^\circ)&\leqslant\underline{\lim_{k\to\infty}}\ \mu_k(A^\circ)\leqslant\underline{\lim_{k\to\infty}}\ \mu_k(A)\leqslant\overline{\lim_{k\to\infty}}\mu_k(A)\\&\leqslant\overline{\lim_{k\to\infty}}\mu_k(\overline{A})\leqslant\mu(\overline{A})=\mu(A).\end{aligned}$$

此即

$$\lim_{k\to\infty}\mu_k(A)=\mu(A).$$

最后, v)⇒i) 是显然的. 于是定理获证. □

定理 2 中 i)–v) 都可用作距离空间测度弱收敛的定义. 很多文献上采用第二种形式.

为了应用, 我们对定理 1 作如下的推广.

定理 3　设 $\mu_k\xrightarrow{w}\mu, C$是$R^{(n)}$ 中的闭集且 $\mu(C)=0.g$ 是 $R^{(n)}$ 上有界可测函数且在 C^c 上连续, 则

$$\int gd\mu_k\to\int gd\mu.$$

证　按照定理 2 由 iii)⇒iv) 的证明, 对一切正整数 m 取

$$G_m=\left\{x:d(x,C)<\frac{1}{m}\right\}.$$

$$f_m(x)=\frac{d(x,C)}{d(x,C)+d(x,G_m^c)}.$$

则 f_m 在 $R^{(n)}$ 上连续, $0 \leqslant f_m(x) \leqslant 1$ 且

$$f_m(x) = \begin{cases} 0, & x \in C, \\ 1, & x \in G_m^c. \end{cases}$$

由于 $\overline{G}_m \supset \overline{G}_{m+1}, m = 1, 2, \cdots$ 且 $\bigcap_m \overline{G}_m = C$($\overline{G}_m$表示$G_m$的闭包). 故而由 μ 的有限性及测度的连续性有

$$\lim_{m\to\infty} \mu(\overline{G}_m) = \mu(C) = 0.$$

对任给定的 $\varepsilon > 0$, 存在 m_0 使 $\mu(\overline{G}_{m_0}) < \dfrac{\varepsilon}{3M}$, 其中 M 是 g 的上界. 再由定理 2 的 iv) 知

$$\overline{\lim_{k\to\infty}} \mu_k(\overline{G}_{m_0}) \leqslant \mu(\overline{G}_{m_0}),$$

于是存在 k_1, 当 $k \geqslant k_1$ 时 $\mu_k(\overline{G}_{m_0}) \leqslant \dfrac{\varepsilon}{3M}$.

由于 $\mu_k \xrightarrow{\text{w}} \mu, gf_{m_0}$ 在 $R^{(n)}$ 上有界连续, 由定理 1 知

$$\lim_{k\to\infty} \int gf_{m_0} d\mu_k = \int gf_{m_0} d\mu.$$

故存在 k_2, 当 $k \geqslant k_2$ 时 $\left|\int gf_{m_0} d\mu_k - \int gf_{m_0} d\mu\right| < \varepsilon/3$.

于是, 当 $k \geqslant \max\{k_1, k_2\}$ 时

$$\begin{aligned}
\left|\int g d\mu_k - \int g d\mu\right| &\leqslant \left|\int g d\mu_k - \int f_{m_0} g d\mu_k\right| \\
&\quad + \left|\int f_{m_0} g d\mu_k - \int f_{m_0} g d\mu\right| + \left|\int f_{m_0} g d\mu - \int g d\mu\right| \\
&\leqslant \left|\int_{G_{m_0}} (g - f_{m_0} g) d\mu_k\right| + \left|\int g f_{m_0} d\mu_k - \int g f_{m_0} d\mu\right| \\
&\quad + \left|\int_{G_{m_0}} (g - f_{m_0} g) d\mu\right| \\
&\leqslant \int_{G_{m_0}} |g|(1 - f_m) d\mu_k + \varepsilon/3 + \int_{G_{m_0}} |g|(1 - f_m) d\mu \\
&\leqslant M\mu_k(G_{m_0}) + \varepsilon/3 + M\mu(G_{m_0}) < \varepsilon.
\end{aligned}$$

因此定理获证. □

6.3.2　弱紧致性

一个测度序列是否存在弱收敛的子序列? 回答是否定的. 但是如果对弱收敛去掉 $\mu_k(R^{(n)}) \to \mu(R^{(n)})$ 的要求, 则有下面的定理 4. 它在证明特征函数极限定理时起着重要的作用.

引理 2　设 D 为 $R^{(n)}$ 中的一个稠密子集, $F_D(x)$ 是定义在 D 上的不降函数, 则由等式

$$\lim_{y^{(k)}\uparrow x} F_D(y^{(k)}) = F(x), x \in R^{(n)}, \{y^{(k)}\} \subset D \tag{7}$$

($y^{(k)} \uparrow x$表示$y^{(1)} < y^{(2)} < \cdots$且$y^{(k)} \to x$) 定义的函数 $F(x)$ 是一个不降且左连续的函数.

证　i) 先证由 (7) 定义的 $F(x)$ 是 $R^{(n)}$ 上单值函数. 设 $x \in R^{(n)}$, 由于 D 在 $R^{(n)}$ 中稠密, 故对任一 $\varepsilon > 0$, 有一 $y = (y_1, \cdots, y_n) \in D$ 使

$$x_j - \varepsilon < y_j < x_j, j = 1, \cdots, n.$$

因此可取一序列 $\{y^{(k)}\} \subset D$ 且 $y^{(k)} \uparrow x$. 而由 $F_D(x)$ 的不降性知对于这个序列 $\{y^{(k)}\}$, 极限

$$\lim_{y^{(k)}\uparrow x} F_D(y^{(k)})存在.$$

故欲证 $F(x)$ 在 $R^{(n)}$ 上单值确定, 只需证明对注意 $\{x^{(k)}\} \subset D, \{y^{(k)}\} \subset D$且$x^{(k)} \uparrow x, y^{(k)} \uparrow x$,

$$\lim_{x^{(k)}\to x} F_D(x^{(k)}) = \lim_{y^{(k)}\to x} F_D(y^{(k)}).$$

而欲证此式, 只需证明, 对任何 $y \in D, y < x$,

$$\lim_{x^{(k)}\uparrow x} F_D(x^{(k)}) \geqslant F_D(y), \{x^{(k)}\} \subset D.$$

事实上, 因为 $x^{(k)} \uparrow x, y < x$, 故存在 k_0 使 $y < x^{(k_0)} < x$, 因而由 F_D 的不降性知

$$F_D(y) \leqslant F_D(x^{(k_0)}) \leqslant \lim_{x^{(k)}\uparrow x} F_D(x^{(k)}),$$

故由上述分析知由 (7) 式定义的 $F(x)$ 是 $R^{(n)}$ 上的单值函数.

ii) 证明 $F(x)$ 在 $R^{(n)}$ 上非降. 设 $x, y \in R^{(n)}, x \leqslant y$, 于是有 D 中的序列 $\{x^{(k)}\}, \{y^{(k)}\}$ 存在, 使 $x^{(k)} \uparrow x, y^{(k)} \uparrow y$, 由于$x^{(k)} < x \leqslant y$, 故对每一 k, 有一 h_k 存在, 使

$$x^{(k)} < y^{(h_k)} < y$$

且 $h_1 < h_2 < \cdots$, 于是

$$\begin{aligned} F(x) &= \lim_{x^{(k)} \uparrow x} F_D(x^{(k)}) \leqslant \lim_{y^{(h_k)} \uparrow y} F_D(y^{(h_k)}) \\ &= \lim_{y^{(k)} \uparrow y} F_D(y^{(k)}) = F(y). \end{aligned}$$

iii) 最后往证 $F(x)$ 左连续. 由 $F(x)$ 的定义知对任一 $\varepsilon > 0$, 有一 $y \in D, y < x$ 使

$$F(x) - \varepsilon < F_D(y) \leqslant F(x).$$

又注意到 $F(x)$ 是不降函数, 可知对任何满足 $y < x' \leqslant x$ 的 x' 都有

$$F(x) - \varepsilon < F_D(y) \leqslant F(x') \leqslant F(x).$$

这就证明了 $F(x)$ 的左连续性. 引理获证. □

定理 4 对 $R^{(n)}$ 上任何一致有界 L–S 测度序列 $\{\mu_k\}$ 总存在子序列 $\{\mu_{k'}\}$ 及 L–S 测度 μ, 使对 μ 的一切连续区间 I 有

$$\lim_{k' \to \infty} \mu_{k'}(I) = \mu(I).$$

证 令 D 表示 $R^{(n)}$ 中一切坐标都是有理数的点的集合, 则 D 是可数集. 令 $F_k(x)$ 是与 μ_k 相对应的标准分布函数.

i) 首先证明存在 F_k 的子序列在 D 上收敛向非降函数 $F_D(r)$.

把 D 中点排成序列 $r^{(1)}, r^{(2)}, \cdots$, 考虑数列 $\{F_k(r^{(1)})\}$, 由于它是一个有界数列, 因此由 Bolzano-Weierstrass 引理知它有一收敛子序列 $\{F_{1k}(r^{(1)})\}$, 其次考虑序列 $\{F_{1k}(r^{(2)})\}$, 由同样的理由知它有一收敛子序列 $\{F_{2k}(r^{(2)})\}$, 应用数学归纳法即能证明在 $\{F_k\}$ 中存在着可数个子序列 $\{F_{lk}\}, l = 1, 2, \cdots$ 使得

$$\{F_k\} \supset \{F_{1k}\} \supset \{F_{2k}\} \supset \cdots$$

且

$$\lim_{k \to \infty} F_{lk}(r^{(l)}) \text{存在}, l = 1, 2, \cdots.$$

于是 $\{F_{ll}(x)\}$ 在 D 上收敛①. 令

$$\lim_{l \to \infty} F_{ll}(r) = F_D(r), r \in D.$$

由于每一 $F_{ll}(r)$ 非降且对一切 $r, r' \in D, r \leqslant r', \Delta_{r',r} F_{ll} \geqslant 0$, 因而 $F_D(r)$ 非降且 $\Delta_{r',r} F_D \geqslant 0$.

① 这种技巧在文献上通常称为对角线手续 (diagonal procedure), 是一个常用的技巧.

ii) 根据引理 2, 在 $R^{(n)}$ 上定义的

$$F(x) \hat{=} \lim_{r \uparrow x} F_D(r)$$

是 $R^{(n)}$ 上不降左连续的函数, 且对任何 $a, b \in R^{(n)}, a \leqslant b$, 由于

$$\Delta_{b,a} F = \lim_{\substack{r' \uparrow b \\ r \uparrow a}} \Delta_{r',r} F_D \geqslant 0,$$

因而 $F(x)$ 是 $R^{(n)}$ 上分布函数而且有界. 故存在有限测度 μ 使 F 是 μ 的分布函数.

iii) 称 2^n 个顶点均属于 D 的左闭右开区间为有理区间, 在有理区间类上定义

$$\mu_D([r, r')) = \Delta_{r',r} F_D, r \leqslant r', r, r' \in D.$$

同第 2 章 §2.3 引理 2 一样可证 μ_D 是有限可加的. 今往证对一切有理区间 I 及区间 J, 若 $\overline{I} \subset J^o$, 则$\mu_D(I) \leqslant \mu(J)$;若$I^o \supset \overline{J}$, 则 $\mu_D(I) \geqslant \mu(J)$.

事实上, 若 $I = [r, r'), J = [a, b)$, 由于$\overline{I} \subset J^o$必有$a < r \leqslant r' < b$, 于是存在有理点序列 $\{r_1^{(k)}\}, \{r_2^{(k)}\}$, 对每一正整数 k,

$$r_1^{(k)} < a < r \leqslant r' < r_2^{(k)} < b$$

且 $r_1^{(k)} \uparrow a, r_2^{(k)} \uparrow b$. 因而

$$\Delta_{r_2^{(k)}, r_1^{(k)}} F_D \geqslant \Delta_{r',r} F_D = \mu_D(I).$$

令 $k \to \infty$, 由 F 的定义知

$$\mu(J) = \mu([a, b)) = \lim_{k \to \infty} \Delta_{r_2^{(k)}, r_1^{(k)}} F_D \geqslant \mu_D(I).$$

同理可证: 若 $I^o \supset \overline{J}$, 则$\mu_D(I) \geqslant \mu(J)$.

iv) 最后证明对 μ 的任一连续区间 J,

$$\lim_{k \to \infty} \mu_{kk}(J) = \mu(J).$$

事实上, 任给区间 J, 存在有理区间序列 $\{I_l\}$, 使

$$I_l^o \supset \overline{I}_{l+1}, I_l^o \supset J, l = 1, 2, \cdots$$

且

$$\bigcap_l I_l = \overline{J}.$$

由 iii) 知对任何正整数 l,

$$\mu(I_l) \geqslant \mu_D(I_{l+1}) = \lim_{k\to\infty} \mu_{kk}(I_{l+1}) \geqslant \overline{\lim_{k\to\infty}} \mu_{kk}(J).$$

令 $l \to \infty$ 即有

$$\mu(\overline{J}) \geqslant \overline{\lim_{k\to\infty}} \mu_{kk}(J).$$

同理可证

$$\mu(J^o) \leqslant \underline{\lim_{k\to\infty}} \mu_{kk}(J).$$

若 J 是 μ 的连续区间, 则

$$\mu(\overline{J}) = \mu(J) = \mu(J^o).$$

故有

$$\lim_{k\to\infty} \mu_{kk}(J) = \mu(J). \qquad \square$$

习题及补充

1. 设 A 是 $R^{(n)}$ 中的开集, 试证 A 是以 A 中有理点为中心、有理数为半径的可数个开球之并 (因而 $R^{(n)}$ 中的开集是 Borel 集).

2. 设 $A \subset R^{(n)}, d(x, A) = \inf\limits_{y\in A} d(x, y)$. 试证 $d(x, A)$ 作为 x 的函数在 $R^{(n)}$ 上一致连续.

3. 设 $\mu_k, k = 1, 2, \cdots, \mu$ 是 $R^{(n)}$ 上 L–S 测度, 试证 $\mu_k \xrightarrow{w} \mu$ 的充分必要条件是对 μ 的一切连续区间 $I, \mu_k(I) \to \mu(I)$ 且对任何 $\varepsilon > 0$, 存在 $a, b \in R^{(n)}$ 使得对一切正整数 $k, \mu_k([a, b)^c) < \varepsilon$.

4. 设 F 是 $R^{(n)}$ 上 L–S 测度 μ 的标准分布函数, 则在任何一个与所有坐标轴不垂直的直线上 F 至多有可数多个不连续点.

5. 设 $F_k, F, k = 1, 2, \cdots$ 是 $R^{(n)}$ 中有界标准分布函数, 如果对 $F(x)$ 的一切连续点 $F_k(x) \to F(x)$ 且 $F_k(+\infty) \to F(+\infty)$, 则称 $F_k \xrightarrow{C} F$. 试证 $F_k \xrightarrow{C} F$ 的充分必要条件是 $F_k(+\infty) \to F(+\infty)$ 且在 $R^{(n)}$ 的一个稠密子集上 $F_k(x) \to F(x)$.

6. 设 μ 为 $R^{(n)}$ 中的 L–S 测度, 称 $A \in \mathscr{B}^{(n)}$ 为 μ 连续集, 若 $\mu(\partial A) = 0$. 试证: 一切 μ 连续集构成集代数.

7. 设 μ 是 $R^{(n)}$ 上有限 L–S 测度, $\mathscr{U}$ 是 μ 的一切连续区间作成的类, 试证 $\mathscr{U}$ 具有以下性质:

i) $\mathscr{U}$ 对有限交封闭;

ii) $R^{(n)}$ 中的任意开集是 $\mathscr{U}$ 中有限或可列个集之并.

再者, 若 $\mathscr{U}$ 是任一包含 $R^{(n)}$ 的具有上述性质的 $\mathscr{B}^{(n)}$ 的子集类, 且对一切 $A \in \mathscr{U}, \mu_k(A) \to \mu(A)$, 则$\mu_k \xrightarrow{w} \mu$($\{\mu_k\}$是$R^{(n)}$上一致有界L–S 测度序列).

8. 设 μ_k, μ 是 $R^{(1)}$ 上的 L–S 测度, $\mu_k \xrightarrow{w} \mu$, $g(x)$是$(-\infty, a]$ 上的有界连续函数, $\mu(\{a\}) = 0$, 试证

$$\int_{(-\infty,a)} g(x) d\mu_k \to \int_{(-\infty,a)} g(x) d\mu, (k \to \infty).$$

9. 设 F_k, F 分别是 μ_k, μ 的标准分布函数, $k = 1, 2, \cdots$ 试证: $\mu_k \xrightarrow{w} \mu$ 的充分必要条件是 $F_k \xrightarrow{C} F$.

§6.4　特征函数极限定理

在引言中我们述及: 不仅特征函数与分布律之间存在一一对应, 而且在特征函数的极限函数与分布律的某种收敛意义下的极限分布之间也存在着对应关系. 这一节我们就着手解决这个问题. 本节的讨论仍然是对有界 L–S 测度及其特征函数进行的.

下面的定理是 §6.3 定理 1 的特殊情形.

定理 1　设 $\{\mu_k\}$ 是 $R^{(n)}$ 中 L–S 测度序列且 $\mu_k \xrightarrow{w} \mu$, 则 μ_k 的特征函数必收敛向 μ 的特征函数.

现在我们来解决相反的问题. 为此先引入

定义 1　设 f 是有限 L–S 测度 μ 的特征函数, 称 $R^{(n)}$ 上的函数

$$\hat{f}(u_1, \cdots, u_n) = \int_0^{u_1} \cdots \int_0^{u_n} f(t_1, \cdots, t_n) dt_1 \cdots dt_n,$$

$$(u_1, \cdots, u_n) \in R^{(n)} \tag{1}$$

为 μ 的积分特征函数.

由于 $f(t_1, \cdots, t_n)$ 在任何 $(t_1, \cdots, t_n) (\in R^{(n)})$ 点连续, 故对任何 $u = (u_1, \cdots, u_n) \in R^{(n)}$,

$$\frac{\partial^n \hat{f}(u_1, \cdots, u_n)}{\partial u_1, \cdots, \partial u_n} = f(u_1, \cdots, u_n). \tag{2}$$

因而 μ 的特征函数由它的积分特征函数唯一决定. 故 μ 由其积分特征函数唯一决定. 即积分特征函数与有限 L–S 测度也是一一对应的.

引理 1　设 $\hat{f}$ 是有限 L–S 测度的积分特征函数, 则

$$\hat{f}(u_1, \cdots, u_n) = \int_{R^{(n)}} \prod_{k=1}^{n} \frac{e^{iu_k x_k} - 1}{i x_k} d\mu(x_1, \cdots, x_n). \tag{3}$$

证 因为

$$\hat{f}(u_1,\cdots,u_n)=\int_0^{u_1}\cdots\int_0^{n_n}\int_{R^{(n)}}\exp\left\{i\sum_{k=1}^n t_kx_k\right\}\\ \times d\mu(x_1,\cdots,x_n)dt_1\cdots dt_n,$$

其中里层积分是 L–S 积分, 外层积分是 R 积分. 而由于特征函数是 $t_1,\cdots,t_n$ 的连续函数, 故也可以看成 L 积分.①

由于 $\exp\left\{i\sum\limits_{k=1}^n t_kx_k\right\}$ 在可测空间 $(R_1^{(n)}\times R_2^{(n)},\mathscr{B}_1^{(n)}\times\mathscr{B}_2^{(n)},\mu\times\Lambda)$ 的可测集 $R_1^{(n)}\times I_{u_1,\cdots,u_n}$(参看注) 上是可积的, 其中 $R_i^{(n)}=R^{(n)},\mathscr{B}_i^{(n)}=\mathscr{B}^{(n)},i=1,2,\Lambda$ 是 $R^{(n)}$ 上的 L 测度. 因而根据 Fubini 定理交换积分次序, 即得 (3) 式. □

在以下的叙述中, 如无特别声明, $f,\hat{f}$ 分别表示 $R^{(n)}$ 上 L–S 测度 μ 的特征函数和积分特征函数 (或带有足码).

定理 2 设 $\{\mu_k\}$ 是 $R^{(n)}$ 上一致有界 L–S 测度序列, 则以下二条件等价:

i) 存在有限 L–S 测度 μ, 在 μ 的每一连续区间 I 上 $\mu_k(I)\to\mu(I)$.

ii) $\hat{f}_k\to$ 某一 g.

在上述条件成立时 g 即 μ 的积分特征函数, 即 $\hat{f}=g$.

证 a) 若 i) 成立, 由于对每一 l, $\dfrac{e^{iu_lx_l}-1}{ix_l}$ 对给定的 u_1 是有界且连续的. 当 $|x_l|\to\infty$ 时极限是零, 故函数 $\prod\limits_{l=1}^n\dfrac{e^{iu_lx_l}-1}{ix_l}$ 作为 $(x_1,\cdots,x_n)$ 的函数满足 §6.3 定理 1 推论的条件, 于是由引理 1 及上述推论知, 对任一 $(u_1,\cdots,u_n)\in R^{(n)}$, 当 $k\to\infty$ 时

$$\begin{aligned}\hat{f}_k(u_1,\cdots,u_n)&=\int_{R^{(n)}}\prod_{l=1}^n\frac{e^{iulxl}-1}{ix_l}d\mu_k(x_1,\cdots,x_l)\\&\to\int_{R^{(n)}}\prod_{l=1}^n\frac{e^{iu_lx_l}-1}{ix_l}d\mu(x_1,\cdots,x_n)=\hat{f}(u_1,\cdots,u_n)\end{aligned}$$

b) 若 $\hat{f}_k\to g$, 由 §6.3 定理 4 知存在子序列 $\{\mu_{k'}\}$ 及有限 L–S 测度 μ, 使对 μ 的一切连续区间 I 有 $\mu_{k'}(I)\to\mu(I)$.

由假设知 $\hat{f}_{k'}\to g$. 但由 a) 及 $\{\mu_{k'}\}$ 满足条件 i) 知 $\lim\hat{f}_{k'}=\hat{f}$, 故 $\hat{f}=g$.

---

① 确切地说, 积分 $\int_0^{u_1}\cdots\int_0^{u_n}$ 可以看成形如 $(-1)^k\int\cdots\int\limits_{I_{u_1}\cdots u_n}$ 的积分, 其中 k 表示 $u_1,\cdots,u_n$ 中负数的个数, 而

$$I_{u_1,\cdots,u_n}=\{(t_1,\cdots,t_n):t_i\text{在}0\text{与}u_i\text{之间},i=1,\cdots,n\}.$$

在定理以下的证明中都作这样的理解.

以下往证对 μ 的任一连续区间 I

$$\mu_k(I) \to \mu(I).$$

用反证法：若所述论断不成立，必存在一关于 μ 的连续区间 I_0 使得

$$\mu_k(I_0) \not\to \mu(I_0).$$

可选取 $\{\mu_k\}$ 的子序列 $\{\mu_{k''}\}$ 使

$$\mu_{k''}(I_0) \to M_0 \neq \mu(I_0).$$

利用上节定理 4，必存在一个 $\{\mu_{k''}\}$ 的子序列 $\{\mu_{k'''}\}$ 及 μ_0，使得在 μ_0 的一切连续区间 I 上，

$$\mu_{k'''}(I) \to \mu_0(I).$$

这时 $\mu_0 \neq \mu$. 因为若 I_0 不是 μ_0 的连续区间，则显然 $\mu_0 \neq \mu$，若 I_0 是 μ_0 的连续区间，则

$$\mu_0(I_0) = \lim_{k'''\to\infty} \mu_{k'''}(I_0) = \lim_{k''\to\infty} \mu_{k''}(I_0) = M_0 \neq \mu(I_0).$$

所以仍有 $\mu_0 \neq \mu$. 另一方面，由前述证明知 μ_0 的积分特征函数也是 g，即 μ 与 μ_0 有同一积分特征函数. 这与有限 L–S 测度由其积分特征函数唯一决定矛盾. 定理获证.

定理 3　设 $\{\mu_k\}$ 是一致有界 L–S 测度序列. 若 $f_k \to g$, a.e. (对 L 测度)，则存在一有限 L–S 测度 μ，在 μ 的任一连续区间上

$$\mu_k(I) \to \mu(I)$$

且

$$g(t) = f(t) = \int e^{itx} d\mu(x) \text{ a.e. (对 } L \text{ 测度)}.$$

证　因为 $f_k \to g$, a.e., 并且 f_k 是一致有界连续函数，因此 g 是 a.e. 有界可测函数. 根据控制收敛定理 $\hat{f}_k \to \hat{g}$，其中 $\hat{g}$ 由下述积分定义：

$$\hat{g}(u_1, \cdots, u_n) = (-1)^k \int \underset{I_{u_1}, \cdots, u_n}{\cdots} \int g(t_1, \cdots, t_n) dt_1 \cdots dt_n,$$

由定理 2 知存在 L–S 测度 μ，使对一切 μ 的连续区间

$$\mu_k(I) \to \mu(I)$$

且 μ 以 $\hat{g}$ 为其积分特征函数. 即对一切 $(u_1,\cdots,u_n)\in R^{(n)}$

$$\hat{g}(u_1,\cdots,u_n)=\hat{f}(u_1,\cdots,u_n).$$

由于

$$\frac{\partial^n \hat{f}(u_1,\cdots,u_n)}{\partial u_1\cdots\partial u_n}=f(u_1,\cdots,u_n),(u_1,\cdots,u_n)\in R^{(n)},$$

因而 $\dfrac{\partial^n \hat{g}(u_1,\cdots,u_n)}{\partial u_1\cdots\partial u_n}$ 存在且

$$\begin{aligned}&\frac{\partial^n \hat{g}(u_1,\cdots,u_n)}{\partial u_1\cdots\partial u_n}=\frac{\partial^n \hat{f}(u_1,\cdots,u_n)}{\partial u_1\cdots\partial u_n}\\&=f(u_1,\cdots,u_n),(u_1,\cdots,u_n)\in R^{(n)}.\end{aligned}$$

但由 $\hat{g}$ 的定义知

$$\frac{\partial^n \hat{g}(u_1,\cdots,u_n)}{\partial u_1\cdots\partial u_n}=g(u_1,\cdots,u_n),\text{a.e.(对 } L \text{ 测度)}.$$

因而

$$f=g,\text{a.e. (对 } L \text{ 测度)}.$$ □

现在, 我们证明特征函数极限定理.

定理 4 若 $\mu_k\xrightarrow{\text{w}}\mu$, 则 $f_k\to f$; 反之, 若 $\{\mu_k\}$ 一致有界且 $f_k\to g$, 其中 $g(t_1,\cdots,t_n)$ 在 $(0,\cdots,0)$ 处连续, 则有一有限 L–S 测度 μ 存在, 使得 $\mu_k\xrightarrow{\text{w}}\mu$ 且 $f=g$.

注 定理的后一部分也称为特征函数的连续性定理.

证 定理的前一部分即定理 1, 今证后一部分.

设 $\{\mu_k\}$ 有界且 $f_k\to g$, 则由定理 3 知存在一有限 L–S 测度 μ, 在 μ 的任一连续区间 I 上

$$\mu_k(I)\to\mu(I).$$

只要证明 $\mu_k(R^{(n)})\to\mu(R^{(n)})$, 定理即获证. 事实上, 由定理 3 知

$$f=g,\text{a.e. (}L\text{测度)}.$$

因而若 $u_1,\cdots,u_n$ 皆大于零时

$$\hat{f}(u_1,\cdots,u_n)=\int_0^{u_1}\cdots\int_0^{u_n}g(t_1,\cdots,t_n)dt\cdots dt_n$$

由于 f 和 g 均在零点连续, 因而

$$\begin{aligned}\mu(R^{(n)}) &= f(0,\cdots,0)\lim_{\substack{n_i\to 0+\\ i=1,\cdots,n}}\frac{1}{u_1\cdots u_n}\hat{f}(u_1,\cdots,u_n)\\ &= \lim_{\substack{n_i\to 0+\\ i=1,\cdots,n}}\frac{1}{u_1\cdots u_n}\int_0^{u_1}\cdots\int_0^{u_n}g(t_1,\cdots,t_n)dt_1\cdots dt_n\\ &= g(0,\cdots,0)=\lim_{k\to\infty}f_k(0,\cdots,0)=\lim_{k\to\infty}\mu_k(R^{(n)}).\end{aligned}$$

这就完成了证明. □

下面我们举例说明特征函数极限定理在概率论中的简单应用, 更重要的应用在第 8 章中将会见到.

定理 5(Хинчин定理)　设 $\{\xi_n\}$ 是独立同分布随机变量序列, 且它们的数学期望有限, $E\xi_n=a$, 则对任何 $\varepsilon>0$,

$$\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k-a\right|<\varepsilon\right)=1. \tag{4}$$

证　i) 首先注意, 若能证明

$$\eta_n=\frac{1}{n}\sum_{k=1}^{n}(\xi_k-a)$$

的特征函数

$$f_{\eta_n}(t)\to 1, \tag{5}$$

则定理即获证明. 因为由定理 4 知, 若 (5) 成立, 则有一 $R^{(1)}$ 上的 L–S 测度 μ, 使

$$P_{\eta_n}\xrightarrow{\text{w}}\mu,$$

且 μ 的特征函数是恒为 1 的函数. 由特征函数和 L–S 测度相互唯一决定, 显然 μ 是集中在原点的概率测度, 即

$$\mu(R^{(1)})=\mu(\{0\})=1.$$

于是不以原点为端点的任何区间都是 μ 的连续区间. 由 $P_{\eta_n}\xrightarrow{\text{w}}\mu$ 知对任何 $\varepsilon>0$,

$$\begin{aligned}\underline{\lim}_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k-a\right|<\varepsilon\right) &= \underline{\lim}_{n\to\infty}P_{\eta_n}((-\varepsilon,\varepsilon))\\ &\geqslant \lim_{n\to\infty}P_{\eta_n}\left(\left[-\frac{\varepsilon}{2},\varepsilon\right)\right)=\mu\left(\left[-\frac{\varepsilon}{2},\varepsilon\right)\right)=1.\end{aligned}$$

又因为 P 是概率, 故

$$\lim_{n\to\infty}P\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k-a\right|<\varepsilon\right)=1.$$

ii) 往证 (5) 式. 令 $\xi'_n = \xi_n - a, n = 1, 2, \cdots$ 则由假设知 $\{\xi'_n\}$ 为独立同分布且数学期望为零的随机变量序列, 且 $\eta_n = \sum_{k=1}^{n} \frac{\xi'_k}{n}$. 于是, 由 §6.1 性质 3, 4 知

$$f_{\eta_n}(t) = \prod_{k=1}^{n} f_{\frac{1}{n} f_{\xi'_k}}(t) = \prod_{k=1}^{n} f_{\xi'_k}\left(\frac{t}{n}\right) = \left[f\left(\frac{t}{n}\right)\right]^n,$$

其中 $f(t)$ 为 ξ'_k 共同的特征函数. 由于 $E\xi'_k = 0$, 故由 §6.1 性质 6 即得 $f\left(\frac{t}{n}\right) = 1 + o\left(\frac{t}{n}\right)\left(\frac{t}{n} \to o\right)$. 于是当 t 给定时,

$$\begin{aligned}\log f_{\eta_n}(t) &= \log\left(1 + o\left(\frac{t}{n}\right)\right)^n = n\log\left(1 + o\left(\frac{t}{n}\right)\right)\\ &= no\left(\frac{t}{n}\right) \to o(n \to \infty).\end{aligned}$$

故 (5) 式获证. 即 (4) 式成立. □

定理 6 设 $\{\xi^{(k)}\}$ 为 n 维独立同分布的随机变量序列, 其数学期望与相关方阵有限, 且 $E\xi^{(k)} = m$, 而 $\xi^{(k)}$ 的相关方阵 D 是满秩的, 则对任何 $x = (x_1, \cdots, x_n) \in R^{(n)}$,

$$\begin{aligned}&\lim_{N\to\infty} P\left(\frac{1}{\sqrt{N}} \sum_{k=1}^{N} (\xi^{(k)} - m) < x\right)\\ &= \frac{1}{(2\pi)^{\frac{n}{2}} |D|^{\frac{1}{2}}} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} e^{-\frac{1}{2} tD^{-1}t'} dt_1 \cdots dt_n.\end{aligned} \tag{6}$$

证 令 $\eta^{(k)} = \xi^{(k)} - m = (\xi_1^{(k)} - m_1, \cdots, \xi_n^{(k)} - m_n)$, 则 $\eta^{(k)}$ 为独立同分布且数学期望为零的随机变量序列, 而 $\eta^{(k)}$ 的二阶原点矩方阵即 D. 故由 §6.1 性质 3, 4 知 $\frac{1}{\sqrt{N}} \sum_{k=1}^{N} \eta^{(k)}$ 的特征函数

$$f_N(t_1, \cdots, t_n) = \left[f\left(\frac{t_1}{\sqrt{N}}, \cdots, \frac{t_n}{\sqrt{N}}\right)\right]^N,$$

其中 $f(t_1, \cdots, t_n)$ 为 $\eta^{(k)}$ 的共同特征函数. 由于 $E\eta^{(k)} = 0$, 故由 §6.1 性质 7 即得

$$f\left(\frac{t_1}{\sqrt{N}}, \cdots, \frac{t_n}{\sqrt{N}}\right) = 1 - \frac{1}{2N} \sum_{j,l=1}^{n} d_{jl} t_j t_l + o\left(\frac{t_1^2 + \cdots + t_n^2}{N}\right),$$

其中 $d_{jl} = E(\xi_j^{(k)} - m_j)(\xi_l^{(k)} - m_l)$, 于是当 $(t_1, \cdots, t_n)$ 给定且 $N \to \infty$ 时

$$\log f\left(\frac{t_1}{\sqrt{N}}, \cdots, \frac{t_n}{\sqrt{N}}\right) = \frac{-1}{2N} \sum_{j,l=1}^{n} d_{jl} t_j t_l + o\left(\frac{1}{N}\right).$$

故

$$\log f_N(t_1,\cdots,t_n)=-\frac{1}{2}tDt'+o(1)\to-\frac{1}{2}tDt'(N\to\infty).$$

因而

$$f_N(t_1,\cdots,t_n)\to e^{-\frac{1}{2}tDt'},(N\to\infty).$$

但 $e^{-\frac{1}{2}tDt'}$ 是 $N(0,D)$ 的特征函数, 它自然在零点连续, 故由定理 4 知 $\frac{1}{\sqrt{N}}\sum_{k=1}^{N}\eta^{(k)}$ 的分布 (简称作 P_N) 弱收敛向 $N(0,D)$ 的分布 (记作 μ). 而对 $N(0,D)$ 的分布而言, $(-\infty,x)$ 边界的测度为零, 故由 §6.3 定理 2 知,

$$P_N((-\infty,x))\to\mu((-\infty,x)),$$

即对 $x=(x_1,\cdots,x_n)\in R^{(n)}$,

$$P\left(\frac{1}{N}\sum_{k=1}^{n}\xi^{(k)}-m<x\right)\to\frac{1}{(2\pi)^{\frac{n}{2}}|D|^{\frac{1}{2}}}$$

$$\times\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}e^{-\frac{1}{2}tDt'}dt_1\cdots dt_n.$$

此即 (6) 式. □

我们用特征函数极限定理来证明 §6.2 定义 3 的合理性.

定理 7　设 $D=(d_{jk})$ 是一 $n\times n$ 非负定方阵, $m=(m_1,\cdots,m_n)$ 为一 n 维实向量, 则函数

$$f(t_1,\cdots,t_n)=\exp\left\{it\cdot m'-\frac{1}{2}tDt'\right\}$$

是某一 n 维随机变量的特征函数. (这个 n 维随机变量称为按 n 维正态律分布的. 若 $|D|\neq0$, 则称为满秩的, 若 $|D|=0$, 则称为降秩的.)

证　当 $|D|\neq0$ 时, 则 D 为一正定方阵, 由 §6.2 推论 1 知 $f(t)$ 是 $N(m,D)$ 的特征函数. 今往考虑 $|D|=0$ 的情形.

D 是 $n\times n$ 非负定方阵, 即对任何 $(t_1,\cdots,t_n)\in R^{(n)}$ 且 $(t_1,\cdots,t_n)\neq(0,\cdots,0)$ 都有

$$tDt'=\sum_{j,k=1}^{n}d_{jk}t_jt_n\geqslant0.$$

于是对任一 $\varepsilon>0$ 及 $t\in R^{(n)},t\neq(0,\cdots,0)$ 都有

$$t(D+\varepsilon I)t'=tDt'+\varepsilon t\cdot t'>0.$$

其中 I 表示 $n\times n$ 单位方阵, 即 $D+\varepsilon I$ 是 $n\times n$ 正定方阵. 因此

$$f_k(t_1,\cdots,t_n)=\exp\left\{it\cdot m'-\frac{1}{2}t(D+\frac{1}{k}I)t'\right\}$$

是 $N\left(m,D+\frac{1}{k}I\right)$ 的特征函数, 显然当 $k\to\infty$ 时

$$f_k(t)\to f(t),t\in R^{(n)}$$

且 $f(t)$ 在 $(0,0,\cdots,0)$ 处连续. 故由特征函数的极限定理知 $f(t)$ 是某一 $R^{(n)}$ 上 L–S 测度的特征函数且具有正规性, 与此特征函数对应的标准分布函数是概率分布函数. 因而存在随机变量以 $f(t)$ 为其特征函数.

最后, 我们应用特征函数的极限定理及测度弱收敛的理论得出一个在统计中很有用的结果.

定理 8 设 g 是 $R^{(n)}$ 上任一有限可测函数, 且在 $R^{(n)}-D$ 上连续, 其中 D 是闭集且 $P_\xi(D)=0$. ξ, $\xi^{(k)},k=1,2,\cdots$ 为 n 维实随机变量, 若 $f_{\xi^{(k)}}(t)\to f_\xi(t)$, 对一切 $t\in R^{(n)}$ 成立, 则

$$f_{g(\xi^{(k)})}(t)\to f_{g(\xi)}(t),t\in R^{(1)}.$$

证 由 $f_{\xi^{(k)}}\to f_\xi$ 及定理 4 知 $P_{\xi^{(k)}}\xrightarrow{w}P_\xi$. 对任意给定的 $t\in R^{(1)},\exp\{ig(x)\cdot t\}$ 是 $R^{(n)}$ 上有界可测函数且在 $R^{(n)}-D$ 上连续, 应用 6.3.1 定理 3 知

$$f_{g(\xi^{(k)})}(t)=\int_{R^{(n)}}e^{ig(x)t}dP_{\xi^{(k)}}\to\int_{R^{(n)}}e^{ig(x)t}dP_\xi=f_{g(\xi)}(t).\qquad\square$$

举两个例子说明定理 8 的应用.

例 1 设 ξ 是按 $N(0,D)$ 分布的 n 维随机变量且 $|D|\neq0$, 若 $f_{\xi^{(k)}}\to f_\xi$, 则 $\xi^{(k)}D^{-1}\xi^{(k)'}$ 的极限分布是自由度为 n 的 χ^2-分布.

事实上, 由线性代数知存在正交方阵 C 使

$$CDC'=\Lambda=\begin{pmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_n^2\end{pmatrix}.$$

令

$$\Lambda^{\frac{1}{2}}=\begin{pmatrix}\sigma_1 & & 0\\ & \ddots & \\ 0 & & \sigma_n\end{pmatrix},\quad \zeta=\xi C'\Lambda^{-\frac{1}{2}}.$$

则 ζ 的特征函数

$$f_\zeta(t)=f_\xi(t\Lambda^{-\frac{1}{2}}C)=e^{-\frac{1}{2}t\Lambda^{-\frac{1}{2}}CDC'\Lambda^{-\frac{1}{2}}t'}=e^{-\frac{1}{2}tt'}.$$

即 ζ 服从标准正态分布. 令 $\eta=\zeta\cdot\zeta'$, 则 η 是自由度为 n 的 χ- 变量. 但

$$\eta=\zeta\cdot\zeta'=\xi C'\Lambda^{-\frac{1}{2}}\cdot\Lambda^{-\frac{1}{2}}C\xi'=\xi D^{-1}\xi'.$$

因此 $\xi D^{-1}\xi'$ 是自由度为 n 的 χ^2-变量. 根据定理 8, 令

$$g(x)=xD^{-1}x'.$$

则

$$f_{\xi^{(k)}D^{-1}\xi^{(k)\prime}}(t)=f_{g(\xi^{(k)})}(t)\to f_{g(\xi)}(t)=f_{\xi D^{-1}\xi'}(t),$$

即 $\xi^{(k)D^{-1}}\xi^{(k)\prime}$ 的极限分布是自由度为 n 的 χ^2-分布.

例 2　若 ξ 按 $N(0,I_n)$ 分布, 其中 I_n 是 n 阶单位方阵. $\mathscr{L}(\xi^{(k)})\to\mathscr{L}(\xi)$. 设 $\overline{x}=\dfrac{1}{n}\sum\limits_{j=1}^{n}x_j$,

$$h(x)=\begin{cases}\dfrac{\sqrt{n(n-1)}\overline{x}}{\sqrt{\Sigma(x_i-\overline{x})^2}}, & \text{当存在}i\neq j\text{使}x_i\neq x_j\text{时},\\ 0, & \text{当}x_1=x_2=\cdots=x_n\text{时}.\end{cases}$$

则 $\mathscr{L}(h(\xi^{(k)}))\to\mathscr{L}(h(\xi))=\mathscr{L}(t_{n-1})$. 其中 $\mathscr{L}(t_{n-1})$ 是自由度为 $n-1$ 的 t- 分布.

注意到 $x_1=x_2=\cdots=x_n$ 的集合为 L 测度为零的闭集, 因而对测度 P_ξ 而言也是零测集, 应用定理 8 即得所需结论.

习题及补充

1. i) 设 ξ_λ 是按 Poisson 律分布的随机变量且 $E\xi_\lambda=\lambda$. 试证：当 $\lambda\to\infty$ 时, $\dfrac{\xi_\lambda-\lambda}{\sqrt{\lambda}}$ 的分布律收敛于正态律 $N(0,1)$.

ii) 设 ξ_r 是按 Γ- 分布律

$$p_r(x)=\begin{cases}0, & x\leqslant 0,\\ \dfrac{c^r}{\Gamma(r)}x^{r-1}e^{-cx}, & x>0,\end{cases}\quad(c>0)$$

分布的随机变量, 试证当 $r\to\infty$ 时, $\dfrac{c\xi_r-r}{\sqrt{r}}$ 的分布律收敛于正态律. $N(0,1)$.

2. i) 若 L–S 测度的特征函数 $f_k(t)\to f(t),(t\in R^{(n)})$ 且 $f(t)$ 在零点连续, 则上述极限关系在 t 的任一有限区间上一致成立.

ii) 若特征函数 $f_k(t)\to f(t), f(t)$ 在零点连续, $t_k\to t_0$, 则 $f_k(t_k)\to f(t_0), t_0\in R^{(n)}$. (提示：利用特征函数极限定理及上节引理 1 及定理 1 的证明.)

3. 如果 $\varphi(t)$ 是特征函数, $\psi(t)$ 是这样的一个函数：对于某一数列 $\{h_n\}$(在 $n\to\infty$ 时 $h_n\to\infty$)，乘积

$$\varphi(t)\psi(h_nt)=f_n(t)$$

也是特征函数, 则函数 $\psi(t)$ 是特征函数.

4. 将Хинчин定理推广至多维情形. 设 $\xi^{(k)}=(\xi_1^{(k)},\cdots,\xi_m^{(k)})$ 是 m 维独立同分布的随机变量序列. $E(\xi_1^{(k)},\cdots,\xi_m^{(k)})=(a_1,\cdots,a_m)=a$. 则对任给的 $\varepsilon>0$,

$$P\left(d\left(\frac{1}{n}\sum_{k=1}^{n}\xi^{(k)},a\right)<\varepsilon\right)\to 1.$$

(提示：可用多维特征函数直接证明, 也可以利用定理 5 的结果.)

5. 设 $\{\xi_n\}$ 是随机变量序列, 用 $\beta_{2,n}$ 表示 ξ_n 的二阶矩, $n=1,2,\cdots$. 若 $\beta_{2,n}$ 有界, 则存在 $\{\xi_n\}$ 的子序列 $\{\xi_{n_k}\}$ 依分布律收敛. (提示：考虑 $k>\beta_{2,n}=\int_{-\infty}^{\infty}x^2dF_{\xi_n}(x)\geqslant a^2\int_{-\infty}^{-a}dF_{\xi_n}(x)+a^2\times\int_{a}^{+\infty}dF_{\xi_n}(x)$. 从而有

$$\frac{k}{a^2}\geqslant F_{\xi_n}(-a)+1-F_{\xi_n}(a)=1-[F_{\xi_n}(a)-F_{\xi_n}(-a)].$$

取 a 充分大, 可使 $1-[F_n(a)-F_n(-a)]<\varepsilon$ 一致成立.)

6. 设 $\{\xi_n\}$ 为随机变量序列, $\beta_{r,n}$ 表示 ξ_n 的 r 阶矩. 若 $\beta_{r,n}$ 皆有限且对一切 r, $\lim\limits_{n\to\infty}\beta_{r,n}=\beta_r$ 有限, 则

i) 若 $\mathscr{L}(\xi_n)\to\mathscr{L}(\xi),\{\beta_r\}$ 即 $\mathscr{L}(\xi)$ 的矩序列.

ii) 若 $\{\beta_r\}$ 唯一决定一个概率分布函数 $F(x)$, 则

$$F_{\xi_n}\xrightarrow{C}F.$$

7. 设 ξ 的分布律为 $P(S),S\in\mathscr{B}^{(1)},\{S_1,\cdots,S_r\}\subset\mathscr{B}^{(1)}$ 且两两不交, $\sum\limits_{k=1}^{r}S_k=R^{(1)},P(S_j)=p_i>0$. $\{\xi_n\}$ 是与 ξ 同分布的独立随机变量序列. 对任意自然数 n, 设 $\nu_j^{(n)}$ 表示 $\xi_1,\cdots,\xi_n$ 的值属于 S_j 的个数, 则

$$\mathscr{L}\left(\sum_{j=1}^{r}\frac{(\nu_j^{(n)}-np_j)^2}{np_j}\right)\to\mathscr{L}(\chi_{r-1}^2),$$

其中 χ_{r-1}^2 表示自由度为 $r-1$ 的 χ^2-变量. 提示：按以下步骤证明：

i) $\nu^{(n)}=(\nu_1^{(n)},\cdots,\nu_r^{(n)})$ 服从多项分布

$$f_{\nu^{(n)}}(t)=(p_1e^{it_1}+\cdots+p_re^{it_r})^n;$$

ii) 令

$$\eta_j^{(n)} = \frac{\nu_j^{(n)} - np_j}{\sqrt{np_j}} j = 1, \cdots, r.$$

则 $\eta^{(n)} = (\eta_1^{(n)}, \cdots, n_r^{(n)})$ 的特征函数

$$f_{\eta^{(n)}}(t) \to e^{-\frac{1}{2}tDt'} \widehat{=} f_\eta(t),$$

其中

$$D = \begin{pmatrix} 1-p_1 & -\sqrt{p_1p_2} & \cdots & -\sqrt{p_1p_r} \\ \ddots & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots \\ -\sqrt{p_rp_1} & -\sqrt{p_rp_2} & \cdots & 1-p_r \end{pmatrix} = I_r - \begin{pmatrix} \sqrt{p_1} \\ \vdots \\ \sqrt{p_r} \end{pmatrix} (\sqrt{p_1}, \cdots, \sqrt{p_r});$$

iii) 对二次型 tDt', 存在正交变换 $C, u = tC$ 且 $u_r = \sum_{j=1}^{r} t_j\sqrt{p_j}$, 使 $tDt' = \sum_{j=1}^{r} t_j^2 - \left(\sum_{j=1}^{r} t_j\sqrt{p_j}\right)^2 = \sum_{j=1}^{r} u_j^2 - u_r^2 = \sum_{j=1}^{r-1} u_j^2$. 则 $C'DC = \begin{pmatrix} I_{r-1} & 0 \\ 0 & 0 \end{pmatrix}$. 令 $\xi \widehat{=} \eta C$, 则

$$f_\xi(u) = f_\eta(uC') = e^{-\frac{1}{2}uC'DCu'} = e^{-\frac{1}{2}\sum\limits_{j=1}^{r-1} u_j^2}$$

因此 ξ 是按 $N(0, C'DC)$ 分布的 (降秩的正态分布, $\xi\xi'$ 按自由度为 $r-1$ 的 χ^2-分布. 但

$$\eta \cdot \eta' = \xi C'C\xi' = \xi\xi'.$$

因此 $\eta \cdot \eta'$ 按自由度为 $r-1$ 的 χ^2-分布;

iv) 令

$$g(x) = g(x_1, \cdots, x_r) = \sum_{j=1}^{r} x_j^2$$

是 $R^{(r)}$ 上连续函数.

$$g(\eta) = \eta \cdot \eta'.$$

由 $f_\eta^{(n)}(t) \to f_\eta(t)(n \to \infty)$, 利用定理 8 知

$$f_{\eta^{(n)} \cdot \eta^{(n)'}}(t) \to f_{\eta \cdot \eta'}(t),$$

即得所欲证.

8. 设 $\xi^{(k)}$ 是 $n+m$ 维随机变量, 且

$$f_{\xi^{(k)}}(t) \to e^{-\frac{1}{2}t\cdot t'}$$

设 $\eta^{(k)} = (\xi_1^{(k)}, \cdots, \xi_n^{(k)}), \zeta^{(k)} = (\xi_{n+1}^{(k)}, \cdots, \xi_{n+m}^{(k)}), |\xi^{(k)}| \neq 0$ 的概率为 $1, k = 1, 2, \cdots$. 试证 $\dfrac{m\eta^{(k)} \cdot \eta^{(k)'}}{n\zeta^{(k)} \cdot \zeta^{(k)'}}$ 的极限分布是自由度为 n, m 的 F 分布.

§6.5 特征函数的非负定性

一个函数具备哪些性质, 才是某一 L–S 测度的特征函数? 即什么是特征函数的定义性质? 这一节就要解决这个问题.

若 $f(t)$ 是 L–S 测度 μ 的特征函数, 则由 6.1.3 知 $f(t)$ 是 $R^{(n)}$ 上的连续函数且对任何正整数 m 及复数 $\alpha_1, \cdots, \alpha_m$ 以及任何 $t^{(1)}, \cdots, t^{(m)} \in R^{(n)}$,

$$\sum_{j,k=1}^{m} f(t^{(j)} - t^{(k)})\alpha_j \bar{\alpha}_k = \int_{R^{(n)}} \left| \sum_{k=1}^{m} \alpha_k^{e^{it^{(k)}\cdot x'}} \right|^2 d\mu(x) \geqslant 0. \tag{1}$$

这个性质我们叫做 $f(t), t \in R^{(n)}$ 具有**非负定性**. 我们将证明 $R^{(n)}$ 上具有非负定性的连续函数必为某一有限 L–S 测度的特征函数, 特别还有 $f(0) = 1$, 则必为某一 n 维随机变量的特征函数.

为了证明上述命题, 我们给出下述

定义 1 设 $T^{(n)} = \{kc : k = (k_1, \cdots, k_n), k_j \in J, j = 1, \cdots, n\}$, 其中 J 表示一切整数组成的集合, $c > 0$ 为常数. 若 $f(t)$ 是定义在 $T^{(n)}$ 上的一个有限函数, 且对任何 m 及复数 $\alpha_1, \cdots, \alpha_m$ 以及 $t^{(j)} \in T^{(n)}, j = 1, \cdots, m$,

$$\sum_{j,k=1}^{m} f(t^{(j)} - t^{(k)})\alpha_j \bar{\alpha}_k \geqslant 0$$

成立, 则 $f(t), t \in T^{(n)}$ 称为 $T^{(n)}$ 上的一个非负定函数.

由于方阵 $(f(t^{(j)} - t^{(k)}))$ 是 Hermite 非负定的, 故有

引理 1 若 $f(t)$ 是 $T^{(n)}$ 上的非负定函数, 则

$$f(0) \geqslant 0,\ f(-t) = \overline{f(t)},\ |f(t)| \leqslant f(0),\ t \in T^{(n)}. \tag{2}$$

定理 1 设 $T^{(n)} = \{kc : k = (k_1, \cdots, k_n), k_j \in J, j = 1, \cdots, n\}$, 则 $f(t), t \in T^{(n)}$ 是 $T^{(n)}$ 上的非负定函数的充分必要条件是它在 $T^{(n)}$ 上与一特征函数

$$\int_{[-\frac{\pi}{c}, \frac{\pi}{c}]} e^{itx'} d\mu(x)$$

相等, 其中 $\left[-\frac{\pi}{c},\frac{\pi}{c}\right]$ 表示 $R^{(n)}$ 中集 $\left\{(x_1,\cdots,x_n):|x_k|\leqslant\frac{\pi}{c},k=1,\cdots,n\right\}$, μ 是满足条件

$$\mu\left(R^{(n)}-\left[-\frac{\pi}{c},\frac{\pi}{c}\right]\right)=0,\mu(R^{(n)})=f(0)$$

的 L–S 测度.

证　对任一正整数 m 及 $x=(x_1,\cdots,x_n)\in R^{(n)}$,

$$\begin{aligned}
0&\leqslant\frac{1}{m^n}\sum_{(0,\cdots,0)\leqslant j,k\leqslant(m-1,\cdots,m-1)}f(jc-kc)e^{-ic(j-k)\cdot x}\\
&=\frac{1}{m^n}\sum_{j_1=0}^{m-1}\cdots\sum_{j_n=0}^{m-1}\sum_{k_1=0}^{m-1}\cdots\sum_{k_n=0}^{m-1}f((j_1-k_1)c,\\
&\quad\cdots(j_n-k_n)c)e^{-ic\sum\limits_{l=1}^{n}(j_l-k_l)x_l}\\
&=\sum_{r_1=-m}^{m}\cdots\sum_{r_n=-m}^{m}\left(1-\frac{|r_1|}{m}\right)\cdots\left(1-\frac{|r_n|}{m}\right)\\
&\quad\times f(r_1c,\cdots,r_nc)e^{-ic\sum\limits_{l=1}^{n}r_lx_l}\widehat{=}G_m(x_1,\cdots,x_n).
\end{aligned}\tag{3}$$

则 $G_m(x)$ 为 $R^{(n)}$ 上非负有界可测函数. 令

$$p_m(x_1,\cdots,x_n)=\left(\frac{c}{2\pi}\right)^nG_m(x_1,\cdots,x_n)\chi_{[-\frac{\pi}{c},\frac{\pi}{c}]}(x_1,\cdots,x_n),(x_1,\cdots,x_n)\in R^{(n)}$$

及

$$\mu_m(B)=\int_B\cdots\int p_m(x_1,\cdots,x_n)dx_1\cdots dx_n,\ B\in\mathscr{B}^{(n)}.$$

则

$$\begin{aligned}
\mu_m(R^{(n)})&=\mu_m\left(\left[-\frac{\pi}{c},\frac{\pi}{c}\right]\right)\\
&=\left(\frac{c}{2\pi}\right)^n\int_{-\frac{\pi}{c}}^{\frac{\pi}{c}}\cdots\int_{-\frac{\pi}{c}}^{\frac{\pi}{c}}G_m(x_1,\cdots,x_m)dx_1\cdots dx_m\\
&=\sum_{r_1=-m}^{m}\cdots\sum_{r_n=-m}^{m}\left(1-\frac{|r_1|}{m}\right)\cdots\left(1-\frac{|r_n|}{m}\right)f(r_1c,\cdots,r_nc)\\
&\quad\times\prod_{l=1}^{n}\left[\frac{c}{2\pi}\int_{\frac{-\pi}{c}}^{\frac{\pi}{c}}e^{-icr_lx_l}dx_l\right]=f(0,\cdots,0).
\end{aligned}\tag{4}$$

由 §6.3 定理 4, 对测度序列 $\{\mu_m\}$ 存在子序列 $\{\mu_{m'}\}$ 及有限 L–S 测度 μ, 使对 μ 的一切连续区间 I 有

$$\mu_{m'}(I)\to\mu(I).\tag{5}$$

对于任何包含 $\left[-\frac{\pi}{c},\frac{\pi}{c}\right]$ 的连续区间 $[a,b,)$ 有

$$\mu_{m'}([a,b)) = \mu_{m'}(R^{(n)}) = f(0,\cdots,0).$$

因而 $\mu([a,b)) = f(0,\cdots,0)$. 令 $a\to\infty, b\to\infty$, 即得 $\mu(R^{(n)}) = f(0,\cdots,0)$, 亦即 $\mu_{m'}\xrightarrow{w}\mu$. 由特征函数极限定理知当 $m'\to\infty$ 时, $\mu_{m'}$ 的特征函数 $f_{m'}(t)$ 应以 μ 的特征函数为极限, 即

$$\lim_{m'\to\infty} f_{m'}(t) = \int e^{itx'}d\mu.$$

但

$$\begin{aligned}f_m(ck_1,\cdots,ck_n) &= \int e^{ick\cdot x'}d\mu_m(x)\\ &= \left(\frac{c}{2\pi}\right)^n\int_{-\frac{\pi}{c}}^{\frac{\pi}{c}}\cdots\int_{-\frac{\pi}{c}}^{\frac{\pi}{c}} e^{ic(k\cdot x')}G_m(x_1,\cdots,x_n)dx_1\cdots,dx_n\\ &= \left(1-\frac{k_1}{m}\right)\cdots\left(1-\frac{k_n}{m}\right)f(ck_1,\cdots,ck_n), ck\in T^{(n)}.\end{aligned}$$

于是, 对于固定的 $k=(k_1,\cdots,k_n)$, 令 $m\to\infty$. 得

$$\lim_{m\to\infty} f_m(ck_1,\cdots,ck_n) = f(ck_1,\cdots,ck_n).$$

因此

$$f(ck_1,\cdots,ck_n) = \int e^{ick\cdot x'}d\mu(x),\ ck\in T^{(n)}.$$

即 f 与 μ 的特征函数在 $T^{(n)}$ 上相合. □

现在来解决本节开始所提出的问题.

定理 2(Bochner-Хинчин 定理) $R^{(n)}$ 上函数 $f(t)$ 为某一 L–S 测度的特征函数的充分必要条件是 $f(t)$ 为连续的非负定函数.

证 条件的必要性显然, 今往证充分性,

令 $T_m^{(n)} = \left\{\frac{k}{m}: k=(k_1,\cdots,k_n), k_j\in J, j=1,\cdots,n\right\}$, 其中 J 为整数集. 则 $f(t)$ 是 $T_m^{(n)}$ 上的非负定函数 (m 为任意正整数). 故由定理 1 知对每一 m 有一 L–S 测度 μ_m, 使

$$f\left(\frac{k}{m}\right) = \int_{[-m\pi,m\pi]} e^{i\left(\frac{k}{m}\right)\cdot x'}d\mu_m(x), \frac{k}{m}\in T_m^{(n)},$$

其中 $[-m\pi,m\pi]$ 表示集合 $\{(x_1,\cdots,x_n): -m\pi\leqslant x_l\leqslant m\pi, l=1,\cdots,n\}$ 而 μ_m 满足条件;

$$\mu_m(R^{(n)}-[-m\pi,m\pi]) = 0,\ \mu_m(R^{(n)}) = f(0).$$

令

$$f_m(t)=\int_{R^{(n)}} e^{it\cdot x'}d\mu_m(x), t\in R^{(n)}.$$

则

$$f_m\left(\frac{k}{m}\right)=f\left(\frac{k}{m}\right), k=(k_1,\cdots,k_n), k_l\in J, l=1,\cdots,n.$$

对任何实数 t_j 及 m 来说, 有一 k_j(与 m,t_j 有关), 使 $0\leqslant t_j-\frac{k_j}{m}<\frac{1}{m}$, 故

$$\lim_{m\to\infty}\frac{k_j}{m}=t_j, j=1,\cdots,n.$$

令 $k=(k_1,\cdots,k_n)$, 利用 $f(t)$ 的连续性则有

$$f(t)=\lim_{m\to\infty}f\left(\frac{k}{m}\right)=\lim_{m\to\infty}f_m\left(\frac{k}{m}\right).$$

若能证明 $f(t)=\lim\limits_{m\to\infty}f_m(t), t\in R^{(n)}$, 则条件的充分性获证. 由上式知, 只需证明

$$\lim_{m\to\infty}\left(f_m(t)-f_m\left(\frac{k}{m}\right)\right)=0.$$

但由 6.1.3 性质 5′ 知

$$\begin{aligned}
&\left|f_m(t_1,\cdots,t_n)-f_m\left(\frac{k_1}{m},\cdots,\frac{k_n}{m}\right)\right|\\
\leqslant&\sum_{j=0}^{n-1}\left|f_m\left(t_1,\cdots,t_{j+1},\frac{k_j+2}{m},\cdots,\frac{k_n}{m}\right)\right.\\
&\left.-f_m\left(t_1,\cdots,t_j,\frac{k_{j+1}}{m},\cdots,\frac{k_n}{m}\right)\right|\\
\leqslant&\sum_{j=0}^{n-1}\sqrt{2f_m(0)\left[f_m(0)-Rf_m\left(0,\cdots,0,t_{j+1}-\frac{k_{j+1}}{m},0,\cdots,0\right)\right]}.
\end{aligned}$$

而 $f_m(0)=\mu_m(R^{(n)})=f(0)$.

$$\begin{aligned}
&f_m(0)-Rf_m\left(0,\cdots,0,t_{j+1}-\frac{k_{j+1}}{m},0\cdots,0\right)\\
=&\int_{R^{(n)}}\left[1-\cos\left(\left(t_{j+1}-\frac{k_{j+1}}{m}\right)x_{j+1}\right)\right]d\mu_m(x)\\
=&\int_{[-m\pi,m\pi]}\left[1-\cos\left(\left(t_{j+1}-\frac{k_{j+1}}{m}\right)x_{j+1}\right)\right]d\mu_m(x).
\end{aligned}$$

再由 $0\leqslant t_{j+1}-\frac{k_{j+1}}{m}<\frac{1}{m}, -m\pi\leqslant x_{j+1}\leqslant m\pi$ 知

$$\cos\left(\left(t_{j+1}-\frac{k_{j+1}}{m}\right)x_{j+1}\right)\geqslant\cos\frac{1}{m}x_{j+1}.$$

因此

$$f_m(0)-Rf_m\left(0,\cdots,0,t_{j+1}-\frac{k_{j+1}}{m},0\cdots,0\right)\leqslant\int_{[-m\pi,m\pi]}\left(1-\cos\frac{1}{m}x_{j+1}\right)d\mu_m(x)$$
$$=f_m(0)-Rf_m\left(0,\cdots,0,\frac{1}{m},0,\cdots,0\right)$$
$$=f(0)-Rf\left(0,\cdots,0,\frac{1}{m},0,\cdots,0\right).$$

故当 $m\to\infty$ 时,

$$\left|f_m(t)-f_m\left(\frac{k}{m}\right)\right|\leqslant\sum_{j=0}^{n-1}\sqrt{2f(0)\left[f(0)-Rf\left(0,\cdots,0,\underset{(j+1)}{\frac{1}{m}},0,\cdots,0\right)\right]}\to 0.$$

由上面分析定理获证.

习题及补充

1. 若 $f(t)$ 是特征函数, 则 $\varphi(t)=e^{(f(t)-1)\alpha}(\alpha>0)$ 也是特征函数 (提示: 先证当 $n>\alpha$ 时 $1+\dfrac{\alpha(f(t)-1)}{n}$ 是特征函数).

2. 若 $f(t)$ 是特征函数, 则

$$\varphi(t)=\frac{1}{t}\int_0^t f(s)ds$$

也是特征函数.

3. 试证满足下列各等式的连续函数 $f(t)$ 是特征函数:

$$f(t)=f(-t),\ f(t+2a)=f(t),$$
$$f(t)=\frac{a-t}{a},\ 0\leqslant t\leqslant a.$$

4. 称 $\{\xi_t,t\in(-\infty,\infty)\}$ 为广义平稳过程, 如果对任意 $t,s\in R^{(1)},E\xi_t=a,E(\xi_t-a)\overline{(\xi_s-a)}=R(t-s)$, 其中 $R(t),t\in R^{(1)}$ 称为过程的相关函数. 称 $\{\xi_t,t\in(-\infty,\infty)\}$ 为均方连续的, 如果对任意 $t,\lim\limits_{s\to t}E|\xi_t-\xi_s|^2=0$.

试证: 均方连续的广义平稳过程的相关函数必为 $R^{(1)}$ 上某一 L–S 测度的特征函数.

第 7 章　独立随机变量和

人们在长期实践中认识到, 概率接近于 1 的事件是几乎一定发生的, 概率接近于零的事件是很难发生的. 因此人们把概率充分接近于 1 的事件看作是实际上必然发生的, 而把概率充分接近于零的事件看作实际上不可能发生的.

但是, 什么叫概率充分接近于 1(或零), 这要看问题的重要性而定. 例如: 测量两个村庄间的距离, 测量误差大于 10 米的概率等于 0.02, 就可以认为是小概率事件而加以忽略. 但是如果要建一座大水电站, 发生使该水电站破坏的大水灾的概率等于千分之一, 有时也不得不认真考虑.

而且, 所谓实际上不可能发生, 并非绝对不发生. 事实上, 任一正概率事件, 无论这概率如何小, 当在同一条件下进行重复试验, 试验次数充分大时, 这事件至少发生一次的概率可以任意地接近于 1. 只是说在预定的一次试验中, 它实际上不可能发生.

在实际问题及理论问题中, 概率接近于 1(或零) 的事件具有重要意义. 概率论的基本问题之一就是研究概率接近于 1(或零) 的规律. 大数定律就是这种概率论命题之一. 在这些命题中, 最早建立的是

Bernoulli 大数律 (1713): 设事件 A 在每一次独立试验中发生的概率为 p, 若 μ_n 表示 n 次独立试验中 A 发生的次数, 则对于任一正数 ε

$$P\left(\left|\frac{\mu_n}{n}-p\right|\geqslant\varepsilon\right)\to 0,(n\to\infty), \tag{1}$$

或与此等价

$$P\left(\left|\frac{\mu_n}{n}-p\right|<\varepsilon\right)\to 1,(n\to\infty). \tag*{(2)$'$}$$

Bernoulli 大数律的重要性在于, 它正好是频率具有稳定性这一经验事实在理论上的解释.

1866 年, Чебышев推广了 Bernoulli 大数律, 得到

Чебышев **定理**　设 $\xi_1,\xi_2,\cdots$ 是两两独立随机变量序列, ξ_n 有有限的数学期望且 $\{D\xi_n\}$ 一致有界, 则对任一 $\varepsilon>0$,

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k-\frac{1}{n}\sum_{k=1}^{n}E\xi_k\right|\geqslant\varepsilon\right)\to 0(n\to\infty). \tag{2}$$

称满足 (2) 式的随机变量序列 $\{\xi_n\}$**服从大数定律.**

这个定理包含许多重要的特例, 若 $\xi_1, \xi_2, \cdots$ 独立同分布, $E\xi_k = a$, 则 (2) 式成为

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n}\xi_k - a\right| \geqslant \varepsilon\right) \to 0, (n \to \infty). \tag{3}$$

第 6 章 §6.4 定理 5(Хинчин定理) 在 $\{\xi_n\}$ 独立但不假定方差有限的情形下证明了 (3) 式成立. 这一定理说明当 n 充分大时, $\frac{1}{n}\sum_{k=1}^{n}\xi_k$ 以任意接近于 1 的概率与 a 充分接近. 它是普通测量方法的理论根据.

更一般的问题是研究依概率稳定性的问题. 即对于 $\{\xi_n\}$, 存在 $b_n \uparrow \infty$ 及 $\{a_n\}$, 使得

$$P\left(\left|\frac{\xi_1 + \cdots + \xi_n}{b_n} - a_n\right| \geqslant \varepsilon\right) \to 0, (n \to \infty).$$

这时称 $\{\xi_n\}$**依概率 P 稳定.**

1909 年, Borel 发现了一个较 Bernoulli 大数律更为深入的定理, 即

Borel 加强大数律 在 Bernoulli 情况下,

$$P\left(\frac{\mu_n}{n} \to p\right) = 1, (n \to \infty).$$

这是一种几乎处处收敛性, 研究这一类型的规律称为强大数律, 或一般地, 如果存在常数序列 $b_n \uparrow \infty$ 及 $\{a_n\}$ 使得

$$\frac{\xi_1 + \cdots + \xi_n}{b_n} - a_n \to 0, \text{a.e.}$$

则称 $\xi_1, \xi_2, \cdots$ **是强稳定的.**

在这一章, 我们将讨论随机变量和的收敛性和稳定性. 而且主要讨论独立随机变量和的强收敛. 由于许多现象可以看作独立因素累积的结果. 因而它有着广泛的实际背景. 长期以来. 关于独立性的研究或伴随独立性的研究而提出的一些问题, 成为概率论的主要对象之一. 对于独立随机变量和的极限理论, 已经创造了相当完整而系统的方法, 得到了丰富而深入的结果. 我们在此只介绍几个经典而重要的结果. 对于非独立的情况, 近年来也展开了广泛的研究, 我们不在此介绍了.

§7.1 0-1 律

本节包括两部分内容: 0-1 律 —— 由此可以证明独立随机变量序列或是 a.e. 收敛, 或是 a.e. 发散;0-1 判据 —— 给出两个随机变量序列在相差一个零概率集上

同时收敛并且有相同极限的判别条件. 这些结果在进一步研究随机变量和的极限性质时, 具有基本重要性.

本节讨论的方法主要基于一个简单的思想, 即 $\xi_1,\xi_2,\cdots$ 与 $\xi_n,\xi_{n+1},\cdots$ 具有相同的收敛域, 并且收敛于同一极限, 即随机变量序列的收敛性及其极限函数, 决定于它的任意一项以后的尾.

我们知道, 概率场 $(\Omega,\mathscr{A},P)$ 上的随机变量序列 $\{\xi_n\}$ 的收敛域①

$$C=\bigcap_k\bigcup_n\bigcap_r\left\{|\xi_{n+r}-\xi_n|<\frac{1}{k}\right\}$$

是一可测集, 即 $C\in\mathscr{A}$. 当我们再仔细分析时, 不难看出 $C\in\sigma(\xi_1,\xi_2,\cdots)$, 这只要注意到 $\xi_{n+r}-\xi_n$ 是 $\sigma(\xi_1,\xi_2,\cdots)$ 可测的, 且 $\sigma(\xi_1,\xi_2,\cdots)$ 是一 σ-代数, 因而对可数运算封闭即知.

由于 $\xi_1,\xi_2,\cdots$ 与 $\xi_n,\xi_{n+1},\cdots$ 有相同的收敛域, 而后一序列的收敛域属于 $\sigma(\xi_n,\xi_{n+1},\cdots)$, 于是

$$C\in\sigma(\xi_n,\xi_{n+1},\cdots),n=1,2,\cdots.$$

因此,

$$C\in\mathscr{C}\triangleq\bigcap_{n=1}^{\infty}\sigma(\xi_n,\xi_{n+1},\cdots).$$

类似地, $\xi_n,\xi_{n+1},\cdots$ 是关于 $\sigma(\xi_n,\xi_{n+1},\cdots)$ 的可测函数, 因而它的极限函数 (如果存在的话), 设为 ξ, 是关于 $\sigma(\xi_n,\xi_{n+1},\cdots)$ 的可测函数, 由 $\xi_1,\xi_2,\cdots$ 与 $\xi_n,\xi_{n+1},\cdots$ 具有相同的极限函数 ξ, 故 $\xi_1,\xi_2,\cdots$ 的极限 ξ 是关于一切 $\sigma(\xi_n,\xi_{n+1},\cdots),n=1,2,\cdots$ 可测的, 即 ξ 是 $\mathscr{C}$ 可测函数.

这样一来,$\{\xi_n\}$ 的收敛性, 可以期望从 $\mathscr{C}$ 的性质得出. 下面的 0-1 律就是断言: 当 $\xi_1,\xi_2,\cdots$, 是独立随机变量序列时, $\mathscr{C}$ 只包含概率等于零或 1 的事件. 因而 $\{\xi_n\}$ 的收敛域的概率等于 1 或零, 即 $\{\xi_n\}$ 或是 a.e. 收敛, 或是 a.e. 发散, 并且当 $\{\xi_n\}$a.e. 收敛时, 它的极限函数 ξ 关于 $\mathscr{C}$ 可测, 因而 a.e. 等于一常数.

利用同样的思想可以讨论级数 $\sum\xi_n$ 及序列 $\left\{\dfrac{1}{b_n}\sum\limits_{k=1}^{n}\xi_k\right\},(b_n\to\infty)$, 而得出一定的有用结果.

现在将上面的讨论总结起来, 确切叙述如下: (以下均设 $\{\xi_n\}$ 是 $(\Omega,\mathscr{A},P)$ 上的随机变量序列.)

定义 1 称 σ-代数

$$\mathscr{C}=\bigcap_{n=1}^{\infty}\sigma(\xi_n,\xi_{n+1}\cdots)$$

① 我们这里所说的收敛均指收敛到有限数.

为 $\xi_1,\xi_2,\cdots$ 的**尾 σ-代数**,$\mathscr{C}$ 中的事件称为 $\xi_1,\xi_2,\cdots$ 的**尾事件**,Ω 上关于 $\mathscr{C}$ 可测的函数称为 $\xi_1,\xi_2,\cdots$ 的**尾函数**.

由定义容易得出下面的基本事实:

1° $\varliminf\limits_{n\to\infty}\xi_n$ $\varlimsup\limits_{n\to\infty}\xi_n$ 是 $\xi_1,\xi_2,\cdots$ 的尾函数, 特别当 $\lim\limits_{n\to\infty}\xi_n$ 存在时, 它是 $\xi_1,\xi_2,\cdots$ 的尾函数;

证 因为

$$\varliminf_{m\to\infty}\xi_m=\lim_{m\to\infty}\inf_{k\geqslant m}\xi_k=\lim_{m\to\infty}\inf_{k\geqslant m}\xi_{n+k}$$
$$=\varliminf_{m\to\infty}\xi_{n+m},n=1,2,\cdots,$$

而 $\varliminf\limits_{m\to\infty}\xi_{n+m}$ 是关于 $\sigma(\xi_{n+1},\xi_{n+2},\cdots)$ 的可测函数, 因而 $\varliminf\limits_{m\to\infty}\xi_m$ 是 $\xi_1,\xi_2,\cdots$ 的尾函数. 其余的结论同样明显, 不再详述.

2° $\varliminf\limits_{n\to\infty}\dfrac{\xi_1+\cdots+\xi_n}{b_n}$, $\varlimsup\limits_{n\to\infty}\dfrac{\xi_1+\cdots+\xi_n}{b_n}$ (其中 $b_n\to\infty$) 是 $\xi_1,\xi_2,\cdots$ 的尾函数, 因而当 $\lim\limits_{n\to\infty}\dfrac{\xi_1+\cdots+\xi_n}{b_n}$ 存在时, 它是 $\xi_1,\xi_2,\cdots$ 的尾函数.

证 对任意 n 来说, 显然 $\lim\limits_{m\to\infty}\dfrac{\xi_1+\cdots+\xi_n}{b_{n+m}}=0$, 所以,

$$\varliminf_{m\to\infty}\frac{\xi_1+\cdots+\xi_n}{b_m}=\lim_{m\to\infty}\frac{\xi_1+\cdots+\xi_n}{b_{n+m}}+\varliminf_{m\to\infty}\frac{\xi_{n+1}+\cdots+\xi_{n+m}}{b_{n+m}}$$
$$=\varliminf_{m\to\infty}\frac{\xi_{n+1}+\cdots+\xi_{n+m}}{b_{n+m}}.$$

而后者是关于 $\sigma(\xi_{n+1},\xi_{n+2},\cdots)$ 的可测函数 (对任意 n), 故知 $\varliminf\limits_{n\to\infty}\dfrac{\xi_1+\cdots+\xi_n}{b_n}$ 是 $\xi_1,\xi_2,\cdots$ 的尾函数. 其余结论可完全类似地证明.

3° $\xi_1,\xi_2,\cdots$ 的收敛域, $\sum\limits_{k=1}^{\infty}\xi_k$ 的收敛域以及 $\dfrac{\xi_1+\cdots+\xi_n}{b_n}(b_n\to\infty)$ 的收敛域, 都是 $\xi_1,\xi_2,\cdots$ 的尾事件.

证 $\xi_1,\xi_2,\cdots$ 的收敛域属于尾 σ-代数 $\mathscr{C}$ 因而是尾事件, 在本节开头已经阐明. 其余的论断可由 $\sum\limits_{n=1}^{\infty}\xi_n$ 与 $\sum\limits_{m=0}^{\infty}\xi_{n+m}$ 的收敛域相同及 $\dfrac{\xi_1+\cdots+\xi_n}{b_n}$ 与 $\dfrac{\xi_{n+1}+\cdots+\xi_{n+m}}{b_{n+m}}$ $(m\to\infty)$ 的收敛域相同这些明显的事实推知.

定理1(Колмогоров 0-1律) 独立随机变量序列的尾事件的概率是 0 或是 1.

证 首先容易推知: 事件 A 同 A 本身独立, 即 $P(A)=P(A\bigcap A)=P(A)\cdot P(A)$ 的充分必要条件是 $P(A)=0$ 或 1.

由此可见, 为了证明定理, 只需证明独立随机变量序列 $\xi_1,\xi_2,\cdots$ 的尾 σ-代数 $\mathscr{C}$ 同 $\mathscr{C}$ 本身独立即可. (由此,$\mathscr{C}$ 中的每一事件, 即尾事件, 都与自身独立, 因而其概率等于 0 或 1.)

事实上, $\mathscr{C} \subset \sigma(\xi_{n+1}, \xi_{n+2}, \cdots)$(对一切 n), 而 $\sigma(\xi_1, \cdots, \xi_n)$ 与 $\sigma(\xi_{n+1}, \xi_{n+2}, \cdots)$ 独立, 故 $\mathscr{C}$ 与 $\sigma(\xi_1, \cdots, \xi_n)$ 独立 (对一切 n), 因而 $\mathscr{C}$ 与 $\bigcup\limits_{n=1}^{\infty} \sigma(\xi_1, \cdots, \xi_n)$ 独立, 注意到后者对有限交封闭, 由独立类扩张定理 (第一章 §1.6 定理 1) 即得 $\mathscr{C}$ 与 $\sigma\left(\bigcup\limits_{n=1}^{\infty} \sigma(\xi_1, \cdots, \xi_n)\right) = \sigma(\xi_1, \xi_2, \cdots)$ 独立. 但 $\mathscr{C} \subset \sigma(\xi_1, \xi_2, \cdots)$, 故 $\mathscr{C}$ 同 $\mathscr{C}$ 独立. □

推论 1　任何独立随机变量序列的尾函数几乎处处等于一常数.

证　设 ξ 是独立随机变量序列 $\xi_1, \xi_2, \cdots$ 的一个尾函数, 根据 0-1 律, 对于任意实数 x, 由于 $\{\xi < x\}$ 是一尾事件, 因而 $P(\xi < x) = 0$ 或 1.

于是有以下三种可能:

(i) 对于任意实数 x, $P(\xi < x) = 1$, 这时显然 $P(\xi = -\infty) = \lim\limits_{n\to\infty} P(\xi < -n) = 1$, 即 $\xi = -\infty$, a.e.;

(ii) 对于每一实数 x, $P(\xi < x) = 0$, 这时显然 $P(\xi = +\infty) = \lim\limits_{n\to\infty} P(\xi \geqslant n) = 1$, 即 $\xi = +\infty$, a.e.;

(iii) 存在实数 $x_1 < x_2$(有限), 使 $P(\xi < x_1) = 0, P(\xi < x_2) = 1$, 取 $a = \sup\{x : P(\xi < x) = 0\}$, 由假设, a 是一有限实数, 这时对于任意 n,

$$P\left(\xi < a - \frac{1}{n}\right) = 0, P\left(\xi < a + \frac{1}{n}\right) = 1,$$

于是

$$P\left(a - \frac{1}{n} \leqslant \xi < a + \frac{1}{n}\right) = 1.$$

令 $n \to \infty$, 得知 $P(\xi = a) = 1$, 即 $\xi = a$, a.e. □

推论 2　设 $\xi_1, \xi_2, \cdots$ 是独立随机变量序列. 则

i) $\xi_1, \xi_2, \cdots$ 或是 a.e. 收敛, 或是 a.e. 发散. 且当 $\xi_1, \xi_2, \cdots$ a.e. 收敛时, 它的极限函数 a.e. 等于一个有限常数;

ii) $\sum\limits_{n=1}^{\infty} \xi_n$ 或是 a.e. 收敛, 或是 a.e. 发散;

iii) $\dfrac{\xi_1 + \cdots + \xi_n}{b_n}$(其中 $b_n \to \infty$) 或是 a.e. 收敛, 或是 a.e. 发散, 且当 a.e. 收敛时, 它的极限函数 a.e. 等于一有限常数.

证　i) $\xi_1, \xi_2, \cdots$ 的收敛域是 $\xi_1, \xi_2, \cdots$ 的尾事件, 所以它的概率等于 0 或 1, 即 $\xi_1, \xi_2, \cdots$ a.e. 收敛或 a.e. 发散. 当它 a.e. 收敛时, 它的极限函数不能 a.e. 等于 $+\infty$ 或 $-\infty$, 由推论 1 即知 $\xi_1, \xi_2, \cdots$ a.e. 收敛于一个有限常数.

ii) 由 3°, $\sum\limits_{n=1}^{\infty} \xi_n$ 的收敛域是 $\xi_1, \xi_2, \cdots$ 的尾事件, 它的概率等于 0 或 1 .

iii) 同 i) 的证明完全类似. □

现在转到本节的第二部分内容.

在研究随机变量序列的极限性质时, 常常借助于另一与已知序列具有同样极限性质的随机变量序列来得出. 为此需要给出判别两个随机变量序列是否具有同样极限性质的条件. 这里所说的具有相同的极限性质以及所要探求的判别条件也主要基于对"尾"的研究得出. 为了讨论这一问题, 首先引入一记号①:

设 $A_1, A_2, \cdots$ 是一事件序列, 我们记

$$\bigcap_n \bigcup_{k \geqslant n} A_k = \{A_n \text{i.o.}\}.$$

不难理解, $\bigcap\limits_n \bigcup\limits_{k\geqslant n} A_k$ 的实际意义是表示 A_n 中有无穷多个发生, 因此 $\{A_n\text{i.o.}\}$ 读作:"有无穷多个 A_n 发生".

以下设 $\{\xi_n\}$ 与 $\{\xi'_n\}$ 是两个随机变量序列.

定义 2 如果 $P(\xi_n \neq \xi'_n \text{i.o.}) = 0$, 则称 $\{\xi_n\}$ 与 $\{\xi'_n\}$**尾等价,** 简称 ξ_n 与 ξ'_n 尾等价.

由定义 2 容易看出, $\{\xi_n\}$ 与 $\{\xi'_n\}$ 尾等价的意义是: 使得 $\xi_1(\omega), \xi_2(\omega), \cdots$ 与 $\xi'_1(\omega), \xi'_2(\omega), \cdots$ 中有无穷多项不相等的 ω 所成的集合只是一个概率为零的事件. 因此, 对于一个零概率集以外的每一 ω, 都存在一个 $n(\omega)$, 使得当 $n > n(\omega)$ 时, $\xi_n(\omega)$ 与 $\xi'_n(\omega)$ 相等. 这样, 尾等价的两个随机变量序列的收敛域只差一个零概率集. 特别其中之一 a.e. 收敛时, 另一亦必 a.e. 收敛. 此外, 它们在共同的收敛域上除去一个零概率集外具有相同的极限.

如果我们仅注意比较收敛域时, 还有下面的概念.

定义 3 如果 $\{\xi_n\}$ 与 $\{\xi'_n\}$ 的收敛域 C 和 C' 至多相差一个零概率集, 即 $P(C \triangle C') = 0$, 则称 $\{\xi_n\}$ 与 $\{\xi'_n\}$**收敛等价**.

显然尾等价蕴含收敛等价.

下面着手研究尾等价和收敛等价的判别条件.

引理 1(Borel-Contelli) 如果 $\sum\limits_{n=1}^{\infty} P(A_n) < \infty$, 则 $P(A_n\text{i.o.}) = 0$.

证 $P(A_n\text{i.o.}) = P\left(\bigcap\limits_n \bigcup\limits_{k\geqslant n} A_k\right) \leqslant P\left(\bigcup\limits_{k\geqslant n} A_k\right) \leqslant \sum\limits_{k=n}^{\infty} P(A_k)$ 对任意 n 成立, 于是令 $n \to \infty$, 根据 $\sum\limits_{n=1}^{\infty} P(A_n) < \infty$ 即得结论. □

① 关于 $\{A_n\text{i.o.}\}$ 及 Borel 0-1 判据在第一章 §1.6 习题 2 中出现过.i.o. 表示 infinite occur 中两个字的字首.

定理 2(Borel 0-1 判据) 设 $\{A_n\}$ 是独立事件序列, 则

$$P(A_n \text{i.o.}) = \begin{cases} 0, & \text{当} \sum_{n=1}^{\infty} P(A_n) < \infty, \\ 1, & \text{当} \sum_{n=1}^{\infty} P(A_n) = \infty. \end{cases}$$

证 利用引理 1 直接可得前一论断. 现在假定 $\sum_{n=1}^{\infty} P(A_n) = \infty$, 往证 $P(A_n$ i.o.$) = 1$. 首先我们有等式

$$\begin{aligned} P(A_n \text{i.o.}) &= P\left(\bigcap_n \bigcup_{k \geqslant n} A_k\right) \\ &= \lim_{n\to\infty} \lim_{m\to\infty} P\left(\bigcup_{k=n}^{m} A_k\right) \\ &= 1 - \lim_{n\to\infty} \lim_{m\to\infty} \prod_{k=n}^{m} P(A_k^c). \end{aligned}$$

利用熟知不等式

$$e^{-x} \geqslant 1 - x, x \geqslant 0$$

可得

$$\exp\left\{-\sum_{k=n}^{m} P(A_k)\right\} \geqslant \prod_{k=n}^{m} [1 - P(A_k)] = \prod_{k=n}^{m} P(A_k^c).$$

由于 $\sum P(A_n) = \infty$, 所以在上式中令 $m \to \infty$, 得出 $\lim_{m\to\infty} \prod_{k=n}^{m} P(A_k^c) = 0$, 从而证明了 $P(A_n \text{i.o.}) = 1$.

推论 3 如果随机变量 $\xi_1, \xi_2, \cdots$ 独立, 则 $\xi_n \xrightarrow{\text{a.e.}} 0$ 的充分必要条件是对于任意正数 c,

$$\sum_{n=1}^{\infty} P(|\xi_n| \geqslant c) < \infty.$$

证 根据 a.e. 收敛的判别条件,$\xi_n \xrightarrow{\text{a.e.}} 0$ 的充分必要条件是对于任意正数 c,

$$P(|\xi_n| \geqslant c \text{ i.o.}) = P\left(\bigcap_n \bigcup_{k \geqslant n} \{|\xi_k| \geqslant c\}\right) = 0.$$

由于 $\{|\xi_n| \geqslant c\}, n = 1, 2, \cdots$ 是独立事件, 所以根据, 0-1 判据, 当且仅当 $\sum_{n=1}^{\infty} P(|\xi_n| \geqslant c) < \infty$ 时上式成立. □

引理 2(收敛等价引理) 如果 $\sum\limits_n P(\xi_n \neq \xi'_n) < \infty$, 则

i) ξ_n 与 ξ'_n 尾等价;

ii) $\sum \xi_n$ 与 $\sum \xi'_n$ 收敛等价;

iii) $\dfrac{\xi_1+\cdots+\xi_n}{b_n}$ 与 $\dfrac{\xi'_1+\cdots+\xi'_n}{b_n}(b_n\to\infty)$ 收敛等价, 且在它们共同的收敛域上有 a.e. 相同的极限.

证 i) 可由 Borel-Contalli 引理直接推出.

ii) 令 $A=\{\xi_n\neq\xi'_n \text{i.o.}\}$, 由 i) 知 $P(A)=0$, 又设 C 和 C' 分别是 $\sum\xi_n$ 与 $\sum\xi'_n$ 的收敛域, 显然

$$\begin{aligned}\omega\in C\bigcap A^c &\Longleftrightarrow \sum\xi_k(\omega)\text{收敛且存在}n(\omega),\text{使对一切}n>n(\omega),\xi_n(\omega)=\xi'_n(\omega)\\ &\Longleftrightarrow \sum\xi'_k(\omega)\text{收敛且存在}n(\omega),\text{使对一切}n>n(\omega),\xi_n(\omega)=\xi'_n(\omega)\\ &\Longleftrightarrow \omega\in C'\bigcap A^c,\end{aligned}$$

即 $C\bigcap A^c=C'\bigcap A^c$. 所以

$$C\bigtriangleup C'=[(C-C')\bigcap A]\bigcup[(C'-C)\bigcap A]\subset A.$$

故

$$P(C\bigtriangleup C')=0,$$

即 $\sum\xi_n$ 与 $\sum\xi'_n$ 收敛等价.

iii) 的证明与 ii) 类似. 只需注意到本节开头关于基本事实 2° 的证明中的等式, 也容易得出在共同的收敛域与 A^c 的交上有相同的极限.

习题及补充

1. 试证 $\{\xi_n\}$ 服从大数定律的充分必要条件是

$$\lim_{n\to\infty} E\frac{\left(\sum\limits_{k=1}^n(\xi_k-E\xi_k)\right)^2}{n^2+\left(\sum\limits_{k=1}^n(\xi_k-E\xi_k)\right)^2}=0.$$

2. 应用第 1 题证明: 若 $\{\xi_n\}$ 是随机变量序列, 且 $\lim\limits_{n\to\infty}\dfrac{1}{n^2}\times D\left(\sum\limits_{k=1}^n\xi_k\right)=0$, 则 $\{\xi_n\}$ 服从大数定律. (Марков定理).

3. 设 $\xi_n, n=1,2,\cdots$ 有有限的数学期望和有界的方差且两两不相关, 试证 $\{\xi_n\}$ 服从大数定律.

4. 设 $\{\xi_n\}$ 是独立随机变量序列, ξ 的数学期望存在且

$$E(\xi|\xi_n) \xrightarrow{\text{a.e.}} \eta,$$

试证 η a.e. 为一常数.

5. 设 $\{\xi_n\}$ 是独立随机变量序列, 且 $\xi_n \xrightarrow{\text{a.e.}} \xi$ 试证对任何正数 c,

$$\sum_n P(|\xi_n - \xi| \geqslant c) < \infty.$$

§7.2　三级数定理与Колмогоров加强大数律

本节研究独立随机变量和的收敛性与稳定性的几个结果. 由于 a.e. 收敛蕴含依概率收敛, 所以我们主要讨论 a.e. 收敛性和 a.e. 稳定性.

在研究随机变量序列的收敛性时, 估计概率 $P(|\xi_n| \geqslant \varepsilon)$ 具有基本的重要性. 在第 3 章 §3.6 习题 1 中, 我们证明了下述

基本不等式. 设 ξ 是一随机变量, g 是 $R^{(1)}$ 上的一个非负函数.

i) 如果 g 在 $[0,\infty)$ 上非降, 并且 g 是一偶函数, 则对任意 $a \geqslant 0$,

$$\frac{Eg(\xi) - g(a)}{\text{a.e.sup } g(\xi)} \leqslant P(|\xi| \geqslant a) \leqslant \frac{Eg(\xi)}{g(a)}; \tag{1}$$

ii) 如果 g 在 $R^{(1)}$ 上非降, 则对于任意实数 a,

$$\frac{Eg(\xi) - g(a)}{\text{a.e.sup } g(\xi)} \leqslant P(\xi \geqslant a) \leqslant \frac{Eg(\xi)}{g(a)}, \tag{2}$$

其中 a.e. $\sup g(\xi) = \inf\{s : P(g(\xi) \geqslant s) = 0\}$.

在基本不等式中取不同的 $g(x)$, 可以得到不同的具体不等式.

对于独立 (或者不相关) 的随机变量 $\xi_1, \cdots, \xi_n$ 的和 $S_n = \sum\limits_{k=1}^{n} \xi_k$, 当 ξ_k 具有有限的数学期望和方差时, 取 $g(x) = x^2$, 可以用 ξ_k 的方差, $k = 1, \cdots, n$ 给出概率 $P(|S_n - ES_n| \geqslant \varepsilon)$ 的界, 即

$$\frac{\sum\limits_{k=1}^{n} D\xi_k - \varepsilon^2}{\text{a.e.sup}|S_n - ES_n|^2} \leqslant P(|S_n - ES_n| \geqslant \varepsilon) \leqslant \frac{\sum\limits_{k=1}^{n} D\xi_k}{\varepsilon^2}. \tag{3}$$

上述不等式的右端即是著名的Чебыщев不等式, 本章引言中提到的Чебыщев定理, 立即可以从这个不等式得出.

不过, 对于 a.e. 收敛性这个不等式还不够, 下面我们将证明的 Колмогоров 不等式, 给出 $P(\max\limits_{k\leqslant n} |S_k - ES_k| \geqslant \varepsilon)$ 的界, 这个界也是用方差表出的.

在用方差作估计界时, 需要假定给定的随机变量具有有限方差. 这看来对于应用有一定的局限性, 不过这可以用截割法来消除, 同时用方差界时, 代替和 S_n, 我们考虑的是 $S_n - ES_n$, 这又引出所谓中心化的概念.

7.2.1 中心化与截割法

普通将 ξ 换作 $\xi - c$ 时, 我们就说将 ξ **中心化于** c. 特别当 ξ 具有有限数学期望时, 我们常常将 ξ 中心化于它的数学期望 $E\xi$, 这时也简单地说成将ξ **中心化.** ξ 已经中心化的充分必要条件是 $E\xi = 0$. 中心化并不改变随机变量的方差. 事实上, 显然有 $D(\xi - c) = D\xi$.

设 ξ 是一给定的随机变量, c 是一正数, 则

$$\xi^c = \xi x_{|\xi|<c}$$

称为ξ **在** c **处的截割.**

显然, 对任何随机变量, 它在 C 处的截割总具有有限方差 (事实上还具有任何阶矩).

当讨论给定随机变量序列 $\{\xi_n\}$ 的收敛性时, 我们常常取一截割的 $\xi_n^{c_n}$, 通过截割的序列来推断原来序列的极限性质, 这种方法即所谓截割法. 这种方法的采用是基于下述事实: 我们可以选取 c_n 充分大, 使 $P(|\xi_n| \geqslant c_n) < \dfrac{\varepsilon}{2^n}, n = 1, 2, \cdots$. 因而

$$\Sigma P(\xi_n \neq \xi_{n^n}^c) \leqslant \Sigma P(|\xi_n| \geqslant c_n) < \varepsilon.$$

这样一来, 根据收敛等价引理, 我们有

ξ_n 与 $\xi_n^{c_n}$ 尾等价;

$\Sigma\xi_n$ 与 $\Sigma\xi_n^{c_n}$ 收敛等价;

$\dfrac{\xi_1 + \cdots + \xi_n}{b_n}$ 与 $\dfrac{\xi_1^{c_1} + \cdots + \xi_n^{c_n}}{b_n}(b_n \to \infty)$ 收敛等价, 且在共同的收敛域上极限 a.e. 相同.

因此, 利用适当截割的序列可以得出给定序列的极限性质.

我们将本节采用的符号规定如下:

$\xi_1, \xi_2, \cdots$ 是给定概率场 $(\Omega, \mathscr{A}, P)$ 上的独立随机变量序列;

ξ_n^c 表示 ξ_n 在 c 处的截割;

$S_n = \sum\limits_{k=1}^{n} \xi_k, ES_n = \sum\limits_{k=1}^{n} E\xi_k, DS_n = \sum\limits_{k=1}^{n} D\xi_k$ 等等.

7.2.2 Колмогоров 不等式

定理 1(Колмогоров不等式) 设 $\xi_1, \cdots, \xi_n$ 独立, $E\xi_k$ 有限且 $|\xi_k| < c, k = 1, \cdots, n$. 则对于每个 $\varepsilon > 0$,

$$1 - \frac{(\varepsilon+2c)^2}{\sum_{k=1}^{n} D\xi_k} \leqslant P(\max_{k\leqslant n}|S_k - ES_k| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^2}\sum_{k=1}^{n} D\xi_k. \tag{4}$$

如果诸方差之中有一为无穷, 则右半不等式自然成立, 此时左半不等式无意义 (因为必有 $c=\infty$). 因此假定方差都有限, 当 $c=+\infty$ 时, 左半不等式自然成立. 因此我们假定 c 有限而证此不等式.

证　为了简单起见, 不失一般性, 我们假定 ξ_k 都已中心化, 于是所求证的不等式中 $ES_k=0$, 条件中的 $|\xi_k|\leqslant c$ 换为 $|\xi_k|\leqslant 2c$.

令

$$A_k = \{\max_{j\leqslant k}|S_j| < \varepsilon\},$$

$$B_k = A_{k-1} - A_k = \{|S_1|<\varepsilon, \cdots |S_{k-1}|<\varepsilon, |S_k|\geqslant\varepsilon\}.$$

这时 $B_k, k=1,\cdots,n$ 两两不交. 记 $A_o=\Omega$, 则

$$A_n^c = A_o - A_n = (A_0 - A_1) + (A_1 - A_2) + \cdots + (A_{n-1} - A_n) = \sum_{k=1}^{n} B_k$$

且

$$B_k \subset \{|S_{k-1}|<\varepsilon, |S_k|\geqslant\varepsilon\}.$$

i) 由于 $S_k\chi_{B_k}$ 与 $S_n - S_k$ 独立及 $E(S_n - S_k)=0$, 有

$$\int_{B_k}|S_n|^2dP = E|S_n\chi_{B_k}|^2 = E|S_k\chi_{B_k}|^2 + E|(S_n - S_k)\chi_{B_k}|^2 \geqslant E|S_k\chi_{B_k}|^2 \geqslant \varepsilon^2P(B_k).$$

因此,

$$\sum_{k=1}^{n} D\xi_k = E|S_n|^2 \geqslant \int_{A_n^c}|S_n|^2dP = \sum_{k=1}^{n}\int_{B_k}|S_n|^2dP \geqslant \varepsilon^2\sum_{k=1}^{n}P(B_k) = \varepsilon^2P(A_n^c).$$

此即

$$P(\max_{k\leqslant n}|S_k|\geqslant\varepsilon) \leqslant \frac{1}{\varepsilon^2}\sum_{k=1}^{n} D\xi_k.$$

ii) 由于

$$S_{k-1}\chi_{A_{k-1}} + \xi_k\chi_{A_{k-1}} = S_k\chi_{A_{k-1}} = S_k\chi_{A_k} + S_k\chi_{B_k}, \tag{5}$$

注意到 $S_{k-1}, \chi_{A_{k-1}}$ 均与 ξ_k 独立及 $\chi_{A_k}\chi_{B_k}=0$, 对 (5) 式两端的模取平方, 然后取数学期望有

$$E|S_{k-1}\chi_{A_{k-1}}|^2 + D\xi_kP(A_{k-1}) = E|S_k\chi_{A_K}|^2 + E|S_k\chi_{B_k}|^2. \tag{6}$$

再由 $|\xi_k| \leqslant 2c$ 得

$$|S_k\chi_{B_k}| \leqslant |S_{k-1}\chi_{B_k}| + |\xi_k\chi_{B_k}| \leqslant (\varepsilon+2c)\chi_{B_k}. \tag{7}$$

利用 $P(A_{k-1}) \geqslant P(A_n)$ 并将 (7) 式代入 (6) 式可得

$$E|S_{k-1}\chi_{A_{k-1}}|^2 + D\xi_k P(A_n) \leqslant E|S_k\chi_{A_k}|^2 + (\varepsilon+2c)^2 P(B_k).$$

对 $k=1,\cdots,n$ 求和, 最后得到

$$\begin{aligned}\left(\sum_{k=1}^{n} D\xi_k\right)P(A_n) &\leqslant E|S_n\chi_{An}|^2 + (\varepsilon+2c)^2\sum_{k=1}^{n}P(B_k)\\ &\leqslant \varepsilon^2 P(A_n) + (\varepsilon+2c)^2 P(A_n^c) \leqslant (\varepsilon+2c)^2.\end{aligned}$$

稍加变形即是所求的不等式的左半.

$$1-\frac{(\varepsilon+2c)^2}{\sum\limits_{k=1}^{n} D\xi_k} \leqslant P(A_n^c) = P(\max_{k\leqslant n}|S_k| \geqslant \varepsilon).$$ □

7.2.3 收敛性与稳定性

现在来研究独立随机变量和的收敛性. 这里所说的收敛性指的是 a.e. 收敛于随机变量 (有限值可测函数). 下列引理中的 $\{\xi_n\}$ 均为独立随机变量序列.

引理 1 如果 $\sum\limits_{n=1}^{\infty} D\xi_n < \infty$, 则 $\sum\limits_{n=1}^{\infty}(\xi_n - E\xi_n)$a.e. 收敛, 如果 $\sum\limits_{n=1}^{\infty} D\xi_n = \infty$ 且 $|\xi_n| \leqslant c$(有限常数), $n=1,2,\cdots$, 则 $\sum\limits_{n=1}^{\infty}(\xi_n - E\xi_n)$a.e. 发散.

证 $\sum(\xi_n - E\xi_n)$ 的 a.e. 收敛或 a.e. 发散, 就是序列 $S_n - ES_n$ 的 a.e. 收敛或 a.e. 发散. 根据 a.e. 收敛的判别条件, 为了证明定理的前一论断, 只需证明: 对于任一 $\varepsilon > 0$,

$$P\left(\bigcup_k\{|(S_{n+k} - ES_{n+k}) - (S_n - ES_n)| \geqslant \varepsilon\}\right) \to 0, (n\to\infty).$$

事实上, 由于

$$\begin{aligned}&\bigcup_k\{|(S_{n+k} - ES_{n+k}) - (S_n - ES_n)| \geqslant \varepsilon\}\\ =&\bigcup_k\{\max_{\nu\leqslant k}|(S_{n+\nu} - ES_{n+\nu}) - (S_n - ES_n)| \geqslant \varepsilon\},\end{aligned}$$

是一不降集序列的并, 由此并利用 Колмогоров 不等式

$$P\left(\bigcup_k\{|(S_{n+k}-ES_{n+k})-(S_n-ES_n)|\geqslant\varepsilon\}\right)$$
$$=\lim_{k\to\infty}P(\{\max_{\nu\leqslant k}|(S_{n+\nu}-ES_{n+\nu})-(S_n-ES_n)|\geqslant\varepsilon\})$$
$$\leqslant\lim_{k\to\infty}\frac{1}{\varepsilon^2}\sum_{\nu=1}^{k}D\xi_{n+\nu}$$
$$=\frac{1}{\varepsilon^2}\sum_{k=n+1}^{\infty}D\xi_k.$$

由于 $\sum D\xi_k<\infty$, 当 $n\to\infty$ 时上式趋于零, 因此引理的前一论断获证.

为了证明引理的后一论断, 我们来考察 $\sum(\xi_n-E\xi_n)$ 的发散区域:

$$\left\{\sum(\xi_n-E\xi_n)\text{发散}\right\}=\bigcup_{\varepsilon>0}\bigcap_n\bigcup_\nu\{|(S_{n+\nu}-ES_{n+\nu})-(S_n-ES_n)|\geqslant\varepsilon\}$$
$$\supset\bigcap_n\bigcup_\nu\{\max_{k\geqslant\nu}|(S_{n+k}-ES_{n+k})-(S_n-ES_n)|\geqslant\varepsilon\},$$

利用 Колмогоров 不等式及 $|\xi_n|\leqslant c$(有限) 及 $\sum_{k=1}^{\nu}D\xi_{n+k}\to\infty$,(当 $\nu\to\infty$) 得

$$P\left(\bigcup_\nu\{\max_{k\geqslant\nu}|(S_{n+k}-ES_{n+k})-(S_n-ES_n)|\geqslant\varepsilon\}\right)$$
$$=\lim_{\nu\to\infty}P(\max_{k\leqslant\nu}|(S_{n+k}-ES_{n+k})-(S_n-ES_n)|\geqslant\varepsilon)$$
$$\geqslant\lim_{\nu\to\infty}\left[1-\frac{(\varepsilon+2_c)^2}{\sum_{k=1}^{\nu}D\xi_{n+k}}\right]=1.$$

再注意到概率等于 1 的事件的可数交仍具有概率 1, 所以

$$P\left(\sum(\xi_n-E\xi_n)\text{发散}\right)\geqslant P\left(\bigcap_n\bigcup_\nu\{\max_{k\leqslant\nu}|(S_{n+k}-ES_{n+k})-(S_n-ES_n)|\geqslant\varepsilon\}\right)$$
$$=1.$$

即 $\sum(\xi_n-E\xi_n)$ a.e. 发散 □

引理 2　如果 $\{\xi_n\}$ 一致有界且 $\sum\xi_n$ a.e. 收敛, 则 $\sum D\xi_n$ 及 $\sum E\xi_n$ 收敛.

证　考察乘积概率空间 $(\Omega\times\Omega,\mathscr{A}\times\mathscr{A},P\times P)$ 上的随机变量

$$\xi'_n(\omega_1,\omega_2) = \xi_n(\omega_1), n=1,2,\cdots,$$

$$\xi''_n(\omega_1,\omega_2) = \xi_n(\omega_2), n=1,2,\cdots.$$

显然 ξ_n, ξ'_n, ξ''_n 同分布, 且 $\xi'_1, \xi''_1, \xi'_2, \xi''_2, \cdots$ 是一独立随机变量序列, 令

$$\xi^s_n = \xi'_n - \xi''_n, n=1,2,\cdots.$$

显然这也是一个独立随机变量序列, 且对一切正整数 n,

$$|\xi^s_n| \leqslant |\xi'_n| + |\xi''_n| \leqslant 2c,$$

$$E\xi^s_n = E\xi'_n - E\xi''_n = 0,$$

$$D\xi^s_n = D\xi'_n + D\xi''_n = 2D\xi_n.$$

由于 $\Sigma\xi_n$ a.e. 收敛, 显然 $\Sigma\xi'_n, \Sigma\xi''_n$ 也都 a.e. 收敛, 因而 $\Sigma\xi^s_n$ a.e. 收敛. 这样, 利用引理 1, $\Sigma D\xi^s_n$ 收敛. 从而 $\Sigma D\xi_n$ 收敛. 再利用引理 1, $\Sigma(\xi_n - E\xi_n)$a.e. 收敛. 又得 $\Sigma E\xi_n = \Sigma\xi_n - \Sigma(\xi_n - E\xi_n)$ 收敛. □

定理 2(三级数判据) 独立随机变量级数 $\Sigma\xi_n$a.e. 收敛于某一随机变量的充分必要条件是: 对于某一固定的正数 c, 下列三个级数收敛:

i) $\Sigma P(|\xi_n| \geqslant c)$; ii) $\Sigma D\xi^c_n$; iii) $\Sigma E\xi^c_n$.

证 由 i) 的收敛性, 根据收敛等价引理, $\Sigma\xi_n$ 与 $\Sigma\xi^c_n$ 收敛等价; 再由 ii) 的收敛性, 根据引理 1, $\Sigma(\xi^c_n - E\xi^c_n)$ a.e. 收敛; 再由 iii) 的收敛性知 $\Sigma\xi^c_n$ a.e. 收敛. 所以, 当级数 i, ii, iii) 收敛时, $\Sigma\xi_n$ a.e. 收敛.

反之, 假定 $\Sigma\xi_n$ a.e. 收敛, 这时 $\xi_n \xrightarrow{\text{a.e.}} 0$, 利用 0-1 判据的推论, 对于任意正数 c, 级数 i) 收敛, 于是 $\Sigma\xi_n$ 与 $\Sigma\xi^c_n$ 收敛等价, $\Sigma\xi^c_n$ a.e. 收敛, 由此利用引理 2(显然 ξ^c_n 一致有界), 级数 ii) 和 iii) 收敛. □

现在转而讨论独立随机变量序列的稳定性. 为此先证明下面的引理。

引理 3(Toeplitz) 设实数 $a_{nk}(k=1,\cdots,k_n, n=1,2\cdots)$ 满足条件: 对每一固定的 $k, a_{nk} \to 0$(当 $n\to\infty$ 时), 而对每一 $n, \sum\limits_{k=1}^{k_n}|a_{nk}| \leqslant c < \infty$, 令 $x'_n = \sum\limits_{k=1}^{k_n} a_{nk}x_k$, 则

i) 当 $x_n \to 0$ 时 $x'_n \to 0$;

ii) 若 $\sum\limits_{k=1}^{k_n} a_{nk} \to 1, x_n \to x$(有限), 则 $x'_n \to x$;

iii) 特别, 若 $b_n = \sum\limits_{k=1}^{n} a_k \uparrow \infty, x_n \to x$(有限), 则 $\dfrac{1}{b_n}\sum\limits_{k=1}^{n} a_k x_k \to x$.

证 i) 对于任一 $\varepsilon > 0$, 存在一个 n, 使得当 $n \geqslant n_\varepsilon$ 时, $|x_n| < \dfrac{\varepsilon}{c}$, 因而

$$|x'_n| \leqslant \sum_{k<n_\varepsilon} |a_{nk}||x_k| + \varepsilon.$$

令 $n\to\infty$, 再令 $\varepsilon\to 0$, 即得 $x'_n\to 0$.

ii) 可由 i) 直接导出:

$$x'_n=\left(\sum_{k=1}^{k_n}a_{nk}\right)x+\sum_{k=1}^{k_n}a_{nk}(x_n-x)\to x.$$

iii) 取 $a_{nk}=\dfrac{a_k}{b_n}(k=1,\cdots,n)$ 此时不妨设 $\{a_n\}$ 为非负序列, 作为 ii) 的特例即是 iii). □

引理 4(Kronecker)　如果 $\Sigma x_n=S,|S|<\infty,b_n\uparrow\infty$, 则 $\dfrac{1}{b_n}\sum\limits_{k=1}^{n}b_kx_k\to 0(n\to\infty)$.

证　令 $b_0=0,a_k=b_k-b_{k-1}$, 则 $b_n=\sum\limits_{k=1}^{m}a_k$, 再令 $S_{n+1}=\sum\limits_{k=1}^{n}x_k,S_1=0$, 应用引理 3 的 iii),

$$\begin{aligned}\frac{1}{b_n}\sum_{k=1}^{n}b_kx_k&=\frac{1}{b_n}\sum_{k=1}^{n}b_k(S_{k+1}-S_k)\\&=\frac{1}{b_n}\left(\sum_{k=1}^{n}b_kS_{k+1}-\sum_{k=1}^{n}b_kS_k\right)\\&=S_{n+1}-\frac{1}{b_n}\sum_{k=1}^{n}a_kS_k\to S-S=0,(n\to\infty).\end{aligned}$$

□

定理 3　如果独立随机变量序列 $\{\xi_n\}$ 具有有限的数学期望和方差且 $\sum\limits_n\dfrac{D\xi_n}{b_n^2}<\infty$, 其中 $b_n\uparrow\infty$, 则 $\dfrac{S_n-ES_n}{b_n}\xrightarrow{\text{a.e.}}0$.

证　根据假定 $\Sigma D\left(\dfrac{\xi_m}{b_n}\right)=\sum\limits_n\dfrac{D\xi_n}{b_n^2}<\infty$, 利用引理 1,

$$\sum\frac{\xi_n-E\xi_n}{b_n}=\sum\left(\frac{\xi_n}{b_n}-E\frac{\xi_n}{b_n}\right)\text{a.e.收敛}.$$

由此再利用 Kronecker 引理即知 $\dfrac{S_n-ES_n}{b_n}\xrightarrow{\text{a.e.}}0$.

定理 4(Комогоров 加强大数律)　设 $\xi_n,n=1,2,\cdots$ 是与 ξ 同分布的独立随机变量序列, 则 $\dfrac{\xi_1+\cdots+\xi_n}{n}\xrightarrow{\text{a.e.}}$(某一有限)$c$ 的充分必要条件是 $E|\xi|<\infty$, 且当收敛时 $c=E\xi$.

证　设 $A_n=\{|\xi|\geqslant n\},A_o,=\varOmega$, 我们有

$$\begin{aligned}\sum P(A_n)&=\sum(n-1)(P(A_{n-1})-P(A_n))\leqslant\sum E|\xi|\chi_{A_{n-1}-A_n}\\&=E|\xi|\leqslant\sum n(P(A_{n-1})-P(A_n))=1+\sum P(A_n).\end{aligned}$$

即

$$\sum P(A_n) \leqslant E|\xi| \leqslant 1 + \sum P(A_n). \tag{8}$$

若 $\frac{\xi_1 + \cdots + \xi_n}{n} \xrightarrow{\text{a.e.}} c$, 则

$$\frac{\xi_n}{n} = \frac{\xi_1 + \cdots + \xi_n}{n} - \frac{n-1}{n}\frac{\xi_1 + \cdots + \xi_{n-1}}{n-1} \xrightarrow{\text{a.e.}} 0.$$

根据 §1 推论 3 知

$$\sum P(A_n) = \sum P\left\{\frac{|\xi_n|}{n} \geqslant 1\right\} < \infty.$$

于是由 (8) 知 $E|\xi| < \infty$, 条件的必要性获证.

反之, 若 $E|\xi| < \infty$, 今往证 $\frac{S_n}{n} \xrightarrow{\text{a.e.}} E\xi$.

令

$$\begin{aligned} \bar{\xi}_k &= \xi_k x_{\{|\xi_k|<k\}}, \\ \bar{S}_n &= \sum_{k=1}^{n} \bar{\xi}_k. \end{aligned}$$

由 (8) 知

$$\sum P(|\xi_n| \geqslant n) = \sum P(A_n) \leqslant E|\xi| < \infty,$$

因而 $\frac{S_n}{n}$ 与 $\frac{\bar{S}_n}{n}$ 收敛等价且在共同的收敛域上有相同的极限. 所以我们只需证明

$$\frac{\bar{S}_n}{n} \xrightarrow{\text{a.e.}} E\xi.$$

利用控制收敛定理及 ξ_n 与 ξ 同分布知

$$E\bar{\xi}_n = \int_{|\xi|<n} \xi dP \to \int \xi dP = E\xi, (n \to \infty).$$

因而根据 Toeplitz 引理,

$$\frac{E\bar{S}_n}{n} \to E\xi.$$

于是只要证明了 $\frac{\bar{S}_n - E\bar{S}_n}{n} \xrightarrow{\text{a.e.}} 0$ 定理即获证明. 应用定理 3, 需验证 $\sum \frac{D\bar{\xi}_n}{n^2} < \infty$. 事实上,

$$\sum_{n=1}^{\infty} \frac{D\bar{\xi}_n}{n^2} \leqslant \sum_{n=1}^{\infty} \frac{E|\bar{\xi}_n|^2}{n^2} = \sum_{n=1}^{\infty}\sum_{m=1}^{n} \int_{m-1\leqslant|\xi|<m} \frac{|\xi^2|}{n^2} dP$$

$$\leqslant \sum_{n=1}^{\infty}\sum_{m=1}^{n}\frac{m^2}{n^2}P(m-1\leqslant|\xi|<m)$$
$$=\sum_{m=1}^{\infty}\sum_{n=m}^{\infty}\frac{m^2}{n^2}P(m-1\leqslant|\xi|<m)$$
$$=\sum_{m=1}^{\infty}\left(\frac{1}{m^2}+\frac{1}{(m+1)^2}+\frac{1}{(m+2)^2}+\cdots\right)m^2P(m-1\leqslant|\xi|<m)$$
$$\leqslant\sum_{m=1}^{\infty}\left(\frac{1}{m^2}+\frac{1}{m(m+1)}+\frac{1}{(m+1)(m+2)}+\cdots\right)m^2P(m-1\leqslant|\xi|<m)$$
$$=\sum_{m=1}^{\infty}\left(\frac{1}{m^2}+\frac{1}{m}\right)m^2P(m-1\leqslant|\xi|<m)$$
$$=\sum_{m=1}^{\infty}(2+(m-1))P(m-1\leqslant|\xi|<m)$$
$$\leqslant 2\sum_{m=1}^{\infty}P(m-1\leqslant|\xi|<m)$$
$$+\sum_{m=1}^{\infty}\int_{m-1\leqslant|\xi|<m}|\xi|dP$$
$$=2+E|\xi|<\infty.$$
□

习题及补充

1. 推广的 Колмогоров 不等式. 设 $\xi_1,\xi_2,\cdots$ 独立且 $E\xi_n=0,n=1,2,\cdots$. 令 $C=\{\max_{k\leqslant n}|S_k|\geqslant c\}$, 若 $r\geqslant 1$, 则

$$c^rP(C)\leqslant E|S_n|^r\chi_C\leqslant E|S_n|^r.$$

应用此不等式证明：如果 $S_n\xrightarrow{r}S$, 则 $S_n\xrightarrow{\text{a.e.}}S$(其中 S 是有限可测函数).

2. 设 $\theta_n,n=1,2,\cdots$ 独立, 且 $Ee^{i\theta_n}=0,n=1,2,\cdots$. 则级数 $\sum c_ne^{i\theta_n}$ 收敛的充分必要条件是 $\sum|c_n|^2$ 收敛.

3. 设 $P(\xi_n=k)=\dfrac{c}{k^2\log^2k}\left(k\geqslant 2,c^{-1}=\sum\limits_{k=2}^{\infty}\dfrac{1}{k^2\log^2k}\right)$ 并且 $\xi_n,n=1,2,\cdots$ 独立, 试证该序列满足强大数律.

4. 设 ξ_n 的分布列为

$$P(\xi_n=n^\alpha)=P(\xi_n=-n^\alpha)=\frac{1}{2},n=1,2,\cdots.$$

若 $\xi_n,n=1,2,\cdots$ 独立, 试证 $\{\xi_n\}$ 满足强大数律的充分必要条件是 $\alpha<\dfrac{1}{2}$.

5. 若 $\xi_1, \xi_2, \cdots$ 两两不相关, 且有有限的数学期望和方差, $\sum_{n=1}^{\infty} \frac{D\xi_n}{n^2} < \infty$, 试证

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n}(\xi_k - E\xi_k)\right| \geqslant \varepsilon\right) \to 0, (n \to \infty).$$

6. 设 $\xi_n, n = 1, 2, \cdots$ 相互独立且同分布, 有有限的数学期望, 令 $S_n = \xi_1 + \cdots + \xi_n$, $\{\alpha_n\}$ 满足 $\alpha_n = O\left(\frac{1}{n}\right)$, 试证 $\{\alpha_n S_n\}$ 服从强大数定律.

7. 设 $f(x)$ 是区间 $[a, b]$ 上的连续函数, 试用概率方法近似计算积分 $\int_a^b f(x)dx$.

8. 用 η_n 表示无穷独立试验中前 n 次试验事件 A 出现的次数, p 是 A 的概率. 试证: 对任意 $\varepsilon > 0$,

$$P(\overline{\lim_{n\to\infty}}\{\eta_n - np > n\varepsilon\}) = 0,$$

$$P(\overline{\lim_{n\to\infty}}\{\eta_n - np < -n\varepsilon\}) = 0.$$

由此说明: 如果用 x 轴表示 n 的值, 用 y 轴表示 $\eta_n - np$ 的值, 则以概率 1 点 $(n, \eta_n - np)$ 界于二直线 $y = \varepsilon x$ 及 $y = -\varepsilon x$ 之间 (除对有限多个 n 以外). 即是说, 以概率 1,$\eta_n - np$ 只能走出二直线外有限次.(还可以把这两条线改为

$$y = (1+\varepsilon)\sqrt{2xpq\mathrm{loglog}x}, y = -(1+\varepsilon)\sqrt{2xpq\mathrm{loglog}x}$$

参看文献 [3],§3.7.)

第 8 章　中心极限定理

§8.1　问题的提出

在自然科学、工程技术及经济科学中, 我们经常遇到这样一类事件, 这事件是由许许多多彼此不相干的随机因素所决定的, 而这些因素中的每一个都对该事件起不大的影响.

例如, 在工业部门大量生产的过程中成批地制造同一类型的产品. 产品的某一指标与规定的指标之间总会有一定的偏差, 这些偏差是许多独立因素引起的. 例如刀具的振动, 机器运转速度的波动, 操作上的各种原因等等. 如果生产过程是正常的, 这些原因中的每一个也都起不大的影响. 但总的影响却可能使每个产品与标准规格产生偏差.

我们的目的就是要研究这一类事件的规律性. 用概率论的语言来说, 我们要研究一类随机变量的分布规律, 它由许多独立随机变量的和组成, 组成这个和的每一随机变量都非常地“小”. 更确切地说, 我们要研究的是由项数越来越多的独立随机变量的和组成的序列的极限分布律, 即所谓中心极限问题.

第七章研究的大数定律是这种类型的问题之一. 它可以说是研究独立随机变量和收敛到单点分布的规律. 例如

Bernoulli 大数律可等价地叙述为：若 $P(\xi_n=1)=p, P(\xi_n=0)=1-p, \xi_n, n=1,2,\cdots$ 独立, $S_n=\xi_1+\cdots+\xi_n$, 则

$$P\left(\frac{S_n}{n}-p<x\right)\to F(x)=\begin{cases}0, x<0,\\ 1, x>0,\end{cases}$$

即 $\mathscr{L}\left(\sum_{k=1}^{n}\frac{\xi_k-p}{n}\right)\to\mathscr{L}(0)$.

在早期的另一个极限定理是

Laplace 定理　在 Bernoulli 情况下 $(p>0, 1-p>0)$

$$P\left(a\leqslant\frac{S_n-np}{\sqrt{npq}}<b\right)\to\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx, q=1-p.$$

这等价于 $\mathscr{L}\left(\frac{S_n-np}{\sqrt{npq}}\right)\to\mathscr{N}(0,1)$.

这一定理是从 De Moivre 关于二项分布的近似计算公式 $P(S_n = j) \approx \frac{1}{\sqrt{2\pi npq}}e^{-\frac{x^2}{2}}, x = \frac{j-np}{\sqrt{npq}}$, 发展来的.

研究下列随机变量:

$$\begin{gathered}\xi_{11},\\ \xi_{21},\xi_{22},\\ \xi_{31},\xi_{32},\xi_{33},\\ \cdots\cdots\end{gathered}$$

其中 $\xi_{nk}(k=1,\cdots,n)$ 独立, $P(\xi_{nk}=1)=p_n, P(\xi_{nk}=0)=1-p_n=q_n, np_n=\lambda>0$, 就可得到下面的

Poisson 定理 设 $S_n=\sum\limits_{k=1}^{n}\xi_{nk}$, 当 $n\to\infty$ 时

$$P(S_n=k)\to\frac{\lambda^k}{k!}e^{-\lambda}, k=0,1,2,\cdots.$$

这个结果可叙述为和 S_n 的分布律收敛于具有参数 λ 的 Poisson 律 $P(\lambda)$.

由于以上三个定理, 在概率论中引入了三个著名的分布律:

(i) 退化律 (即单点分布)$\mathscr{L}(a)$: 分布函数 $F(x)=\begin{cases}0, x\leqslant a,\\ 1, x>a,\end{cases}$ 特征函数 $f(t)=e^{ita}$;

(ii) 正态律 $N(a,\sigma^2)$: 特征函数 $f(t)=e^{ita-\frac{\sigma^2}{2}t^2}$;

(iii) Poisson 律 $P(\lambda,a,b)$: 特征函数 $f(t)=e^{ita+\lambda(e^{itb}-1)}$.

由于特征函数的引入, 使极限理论的研究有了一个有力的工具, 获得了一系列重要结果. 例如Ляпунов 定理, Lindeberg 条件等对于收敛向正态律得到了很好的结果 (§8.2 详述). 随着对中心极限定理的了解逐步深入, 概括出中心极限问题的一般提法如下:

设 $\xi_{nk}, k=1,2,\cdots,k_n$ 是独立随机变量, $S_n=\sum\limits_{k=1}^{k_n}\xi_{nk}, n=1,2\cdots$.

(1)S_n 的一切可能的极限分布律是哪些分布律?

(2) 求出 S_n 收敛于某一指定的极限分布律的条件.

如果对 ξ_{nk} 不附加任何限制, 那么问题 (1) 就没有什么实际意义, 因为我们可以取一具有任一分布律的随机变量 ξ, 而令

$$\xi_{n1}=\xi, \xi_{nk}=0, k=2,\cdots,k_n.$$

于是

$$\mathscr{L}(S_n) \to \mathscr{L}(\xi).$$

这样, 任何分布律都可以作为某一和 S_n 的极限分布律.

根据我们所提出的问题的实际背景, 今后我们永远假定 $\{\xi_{nk}\}$ 满足**"一致可渐近忽略"条件**:

$$\max_{1\leqslant k\leqslant k_n} P(|\xi_{nk}| \geqslant \varepsilon) \to 0(n\to\infty), \text{对一切}\varepsilon > 0. \tag{U}$$

在条件 (U) 之下, 中心极限问题 (1), (2) 获得完整的解答. 已经证明了一切可能的极限分布律是所谓的无穷可分律. 关于问题 (2) 也找出一般的充分必要条件以及向常见的分布律 —— 退化律、正态律、Poisson 律收敛的条件.

本章我们将在更强的条件下, 即具有有界方差的情形下, 解决中心极限问题. 至于一般的无穷可分律只介绍一些概念和结果. 详细地证明请读者参看有关文献.

§8.2　中心极限定理 —— 具有有界方差情形

现在我们来解决中心极限问题的一个特殊情形, 即所谓具有有界方差的情形. 具体地说, 我们假定独立随机变量 $\xi_{nk}, k=1,\cdots,k_n$ 具有有限方差 $\sigma_{nk}^2 = D\xi_{nk}$, 首先设它们已经中心化, 即 $E\xi_{nk}=0$. 用 $P_{nk}(B), f_{nk}(t)$ 分别表示 ξ_{nk} 的分布律和特征函数, $\prod\limits_{k=1}^{k_n}, \sum\limits_{k=1}^{k_n}$ 和 $\max\limits_{1\leqslant k\leqslant k_n}$ 分别简记作 $\prod\limits_k, \sum\limits_k$ 和 $\max\limits_k$. 极限过程如无特别声明均指对 $n\to\infty$ 来取的. 我们假设下面的条件 (C) 成立:

$$\max_k \sigma_{nk}^2 \to 0, \sum_k \sigma_{nk}^2 \leqslant c < \infty (c\text{与}n\text{无关}). \tag{C}$$

称 (C) 为有界方差条件.

利用Чебышев不等式, 对于任意 $\varepsilon > 0$,

$$\max_k P(|\xi_{nk}| \geqslant \varepsilon) \leqslant \frac{1}{\varepsilon^2}\max_k \sigma_{nk}^2 \to 0.$$

所以 $\{\xi_{nk}\}$ 满足一致可渐近忽略条件. 因而我们现在讨论的是一般中心极限问题的一个特殊情形. 不过就论证的主要步骤来说, 可以反映出一般情形的概况, 只是后者在具体的论证中更为复杂罢了. 本节中的 $\log z$ 一律表示 z 的对数的主值.

引理 1(比较引理)　在条件 (C) 之下, 对于任一固定的 t 来说, 当 $n \geqslant n_t$(充分大) 时 $\log f_{nk}(t)$ 存在 (有限), 且

$$\sum_k \{\log f_{nk}(t) - (f_{nk}(t) - 1)\} \to 0. \tag{1}$$

证 由于 $E\xi_{nk}=0$, 利用第 6 章 §6.1 性质 6 中的公式 (23),(24) 可知

$$f_{nk}(t)=1-\theta_{nk}\frac{\sigma_{nk}^2}{2}t^2,|\theta_{nk}|\leqslant 1.$$

因此根据条件 (C),

$$\max_k|f_{nk}(t)-1|\leqslant\frac{t^2}{2}\max_k\sigma_{nk}^2\to 0,\sum_k|f_{nk}(t)-1|\leqslant\frac{c}{2}t^2.$$

所以存在 n_t, 当 $n\geqslant n_t$ 时, $|f_{nk}(t)-1|\leqslant\frac{1}{2}$, 从而利用熟知的不等式

$$\log(1+z)=z+\theta z^2,|\theta|\leqslant 1,\text{当}|z|\leqslant\frac{1}{2},$$

得知 $\log f_{nk}(t)$ 存在 (有限) 且

$$\log f_{nk}(t)=f_{nk}(t)-1+\theta_{nk}|f_{nk}(t)-1|^2,|\theta_{nk}|\leqslant 1,$$

所以

$$\begin{aligned}\left|\sum_k\{\log f_{nk}(t)-(f_{nk}(t)-1)\}\right|&\leqslant\sum_k|f_{nk}(t)-1|^2\\&\leqslant\max_k|f_{nk}(t)-1|\sum_k|f_{nk}(t)-1|\to 0.\end{aligned}$$

□

我们的目的是要求出 $\sum_k\xi_{nk}$ 的极限分布律. 为此我们只需求出 $\prod_k f_{nk}(t)$ 的极限或等价地求出 $\exp\left\{\sum_k\log f_{nk}(t)\right\}$ 的极限. 由比较引理, 在条件 (C) 之下 $\sum_k\log f_{nk}(t)$ 与 $\sum_k(f_{nk}(t)-1)$ 有相同的极限 (如果极限存在的话). 因此, 为了实现我们的目的, 转而来考虑 $\exp\left\{\sum_k(f_{nk}(t)-1)\right\}$ 的极限. 令

$$\psi_n(t)=\sum_k(f_{nk}(t)-1). \tag{2}$$

由于 $E\xi_{nk}=\int xdP_{nk}=0$, 所以

$$\psi_n(t)=\sum_k\int(e^{itx}-1-itx)\frac{1}{x^2}x^2dP_{nk}=\int\frac{e^{itx}-1-itx}{x^2}x^2d\left(\sum_k P_{nk}\right).$$

令
$$\mu_n(B) = \int_B x^2 d\sum_k P_{nk} = \sum_k \int_B x^2 dP_{nk} = B \in \mathscr{B}^{(1)}. \tag{3}$$
则 μ_n 是 $R^{(1)}$ 上的 L–S 测度, 且一致有界.
$$\mu_n(R^{(1)}) = \int_{R^{(1)}} x^2 d\sum_k P_{nk} = \sum_k \int x^2 dP_{nk} = \sum_k \sigma_{nk}^2 \leqslant c.$$
此外 μ_n 是关于 $\sum_k P_{nk}$ 的不定积分, 而 $\dfrac{d\mu_n}{d\sum_k P_{nk}} = x^2$, 同时
$$\frac{e^{itx}-1-itx}{x^2} \to 0, (|x| \to \infty\text{时}),$$
$$\frac{e^{itx}-1-itx}{x^2} \to -\frac{t^2}{2}, \quad (|x| \to 0).$$
因而 $\dfrac{e^{itx}-1-itx}{x^2}$ 是一有界连续函数, 因此关于 μ_n 可积. 于是根据第 3 章 §3.7 定理 7 推论, 可以将 $\psi_n(t)$ 表成如下的形式:
$$\psi_n(t) = \int \frac{e^{itx}-1-itx}{x^2} d\mu_n. \tag{4}$$
下面我们证明, 当 $e^{\psi_n(t)}$ 的极限 $(n \to \infty)$ 存在时, $e^{\psi_n(t)}$ 的极限具有形式 $e^{\psi(t)}$,
$$\psi(t) = \int \frac{e^{itx}-1-itx}{x^2} d\mu, \tag{5}$$
其中 μ 是有限 L–S 测度, 且对 μ 的任一连续区间 $I, \mu_n(I) \to \mu(I)$, 且 $\mu(R^{(1)}) \leqslant c$. 从而得知满足 (C) 的和的极限分布的特征函数具有形式 $e^{\psi(t)}$. 为此先证明

引理 2　(5) 中的 μ 与 ψ 相互唯一决定.

证　ψ 由 μ 唯一决定是显然的. 今证 μ 由 ψ 唯一决定: 对 $\psi(t)$ 求导, 根据控制收敛定理的推论:
$$\psi'(t) = i\int \frac{e^{itx}-1}{x} d\mu, \tag{6}$$
$$\psi''(t) = -\int e^{itx} d\mu. \tag{7}$$
因此 $-\psi''(t)$ 是有限 L–S 测度的特征函数, 因而由第 6 章 §6.2 知 μ 由 $-\psi''(t)$ 唯一决定, 从而 μ 由 ψ 唯一决定.

引理 3　$e^{\psi_n(t)} \to$(某一)$g(t), t \in R^{(1)}$ 的充分必要条件是存在有限 L-S 测度 μ, 在 μ 的一切连续区间 I 上有 $\mu_n(I) \to \mu(I)$, 且在此条件成立时 $g(t) = e^{\psi(t)}$, 其中 $\psi(t)$ 由 (5) 式确定.

证 由于 $\dfrac{e^{itx}-1-itx}{x^2}$ 是 $R^{(1)}$ 上的有界连续函数且当 $x\to\infty$ 时以 0 为极限, 因此由第 6 章 §6.3 定理 1 推论知, 若对 μ 的一切连续区间 I 有 $\mu_m(I)\to\mu(I)$, 则

$$\int \frac{e^{itx}-1-itx}{x^2}d\mu_n \to \int \frac{e^{itx}-1-itx}{x^2}d\mu.$$

故条件的充分性获证. 并且 $g(t)=e^{\psi(t)}$

以下往证条件的必要性: 设 $e^{\psi_n(t)}\to g(t), t\in R^{(1)}$. 由于 μ_n 一致有界, 根据第 6 章 §6.3.2 定理 4, 存在子序列 $\{\mu_{n'}\}$ 及 μ, 使对 μ 的一切连续区间

$$\mu_{n'}(I)\to\mu(I), n'\to\infty.$$

再由以上所证结果知 $e^{\psi_{n'}(t)}\to e^{\psi(t)}=g(t), \psi(t)$ 由 μ 确定. 与第 6 章 §6.4 证明定理 2 的方法一样可证对 $\{\mu_n\}$ 的任一子序列 $\{\mu_{n''}\}$ 皆有

$$\mu_{n''}(I)\to\mu(I), n''\to\infty.$$

即

$$\mu_n(I)\to\mu(I), n\to\infty. \qquad \square$$

这样, 我们已经证明了满足 (C) 的和的极限分布的特征函数皆有 $e^{\psi(t)}$ 的形式, 其中 $\psi(t)$ 由 (5) 决定. 同时也证明了给定的和的特征函数收敛于指定的极限特征函数 $e^{\psi(t)}$ 的充分必要条件如引理 3 所述. 这里 $g(t)$ 的连续性不需要假定, 而是自然的结果. 即在条件 (C) 之下, 若 $e^{\psi_n(t)}$ 收敛向 $g(t)$, 则 $g(t)$ 必然连续, 因而 $e^{\psi_n(t)}\to e^{\psi(t)}, e^{\psi(t)}$ 必然为特征函数.

引理 4 每一形如 $e^{\psi(t)}$ 的函数 (其中 $\psi(t)$ 如 (5) 所述) 都是一随机变量的特征函数, 具数学期望零和有限方差 $\sigma^2=\mu(R^{(1)})$, 并且是某一满足 (C) 的和的极限分布律的特征函数.

证 首先, 由于

$$\psi(t)=\int \frac{e^{itx}-1-itx}{x^2}d\mu=\lim_{N\to\infty}\int_{[-N,N)}\frac{e^{itx}-1-itx}{x^2}d\mu$$

其被积函数对 x 来说是 $R^{(1)}$ 上有界连续函数, 因而可以通过将 $[-N,N)$ 分成小区间, 用阶梯函数的积分来逼近: 设分法 T 为

$$-N=x_{no}<x_{n1}<\cdots<x_{nk_n}=N, (0\text{不作分点}).$$

令

$$g_T(x)=\sum_{k=0}^{k_n-1}\frac{e^{itx_{nk}}-1-itx_{nk}}{x_{nk}^2}\chi_{[x_{nk},x_{n(k+1)})}(x).$$

于是若令 $l(T)$ 表示分法 T 的最大子区间长度, 则

$$\int_{-N}^{N}\frac{e^{itx}-1-itx}{x^2}d\mu=\lim_{l(T)\to 0}\int_{[-N,N)}g_T(x)d\mu$$
$$=\lim_{l(T)\to 0}\sum_{k=0}^{k_n-1}[ita_{nk}+\lambda_{nk}(e^{itb_{nk}}-1)],$$

其中 $\lambda_{nk}=\frac{1}{x_{nk}^2}\mu([x_{nk},x_{n,(k+1)})),a_{nk}=-\lambda_{nk}x_{nk},b_{nk}=x_{nk}.$ 由此可见

$$\exp\left[\int_{(-N,N)}\frac{e^{itx}-1-itx}{x^2}d\mu\right]$$

是 Poisson 型特征函数的乘积的极限, 且易见它在 $t=0$ 点连续, 因而是特征函数, 因此 $e^{\psi(t)}$ 也是特征函数的极限, 且在 $t=0$ 点连续, 因而 $e^{\psi(t)}$ 是一随机变量的特征函数.

其次,

$$\frac{d}{dt}e^{\psi(t)}=ie^{\psi(t)}\int\frac{e^{itx}-1}{x}d\mu,$$
$$\frac{d^2}{dt^2}e^{\psi(t)}=i^2e^{\psi(t)}\left[\int\frac{e^{itx}-1}{x}d\mu\right]^2+i^2e^{\psi(t)}\int e^{itx}d\mu.$$

利用特征函数的性质得

$$\alpha_1=i^{-1}\frac{d}{dt}e^{\psi(t)}|_{t=0}=0,$$
$$\alpha_2=i^{-2}\frac{d^2}{dt^2}e^{\psi(t)}|_{t=0}=\int d\mu=\mu(R^{(1)}).$$

这里 α_1,α_2 是对应于特征函数 $e^{\psi(t)}$ 的分布的一阶和二阶原点矩. 由于 $\alpha_1=0,\alpha_2$ 也就是相应的方差.

最后, 取同分布独立随机变量 $\xi_{nk},k=1,\cdots,n$ 分别具有特征函数 $e^{\frac{1}{n}\psi(t)}$, 由于

$$\frac{1}{n}\psi(t)=\int\frac{e^{itx}-1-itx}{x^2}d\left(\frac{1}{n}\mu\right),$$

根据本引理的第一部分, $e^{\frac{1}{n}\psi(t)}$ 是随机变量的特征函数, 再由本引理的第二部分, $E\xi_{nk}=0,\sigma_{nk}^2=D\xi_{nk}=\frac{1}{n}\mu(R^{(1)})$, 所以

$$\max_k\sigma_{nk}^2\to 0,\quad \sum_k\sigma_{nk}^2=\mu(R^{(1)})<\infty,$$

即 $\{\xi_{nk}\}$ 满足条件 (C), 而 S_n 的特征函数为

$$\prod_k f_{nk}(t) = e^{\psi(t)} \to e^{\psi(t)}.$$

这样又证明了 $e^{\psi(t)}$ 都是某一满足条件 (C) 的独立随机变量和的极限分布律的特征函数.

定理 1 设 $\xi_{nk}, k=1,\cdots,k_n$ 是独立随机变量. $E\xi_{nk}=0, \sigma_{nk}^2 = D\xi_{nk}$ 有限, 且满足

$$\max_k \sigma_{nk}^2 \to 0, \sum_k \sigma_{nk}^2 \leqslant c < \infty. \tag{C}$$

i) 如果 $\prod\limits_k f_{nk}(t) \to g(t)$, 则 $g(t)$ 必是特征函数且形如 $e^{\psi(t)}$, 其中 $\psi(t)$ 由 (5) 定义. 因此形如 $\sum\limits_k \xi_{nk}$ 的一切可能的极限分布律的特征函数类与一切形如 $e^{\psi(t)}$ 的函数类重合.

ii) $\mathscr{L}\left(\sum\limits_k \xi_{nk}\right) \to \mathscr{L}(\xi)$, 且 $f_\xi(t) = e^{\psi(t)}$(由 i) $f_\xi(t)$ 必具此形式) 的充分必要条件是在 μ 的一切连续区间 I 上 $\mu_n(I) \to \mu(I)$, 其中

$$\mu_n(B) = \sum_k \int_B x^2 dP_{\xi_{nk}}, \tag{8}$$

而 μ 与 $\psi(t)$ 由 (5) 式相互唯一决定.

iii) $\mathscr{L}\left(\sum\limits_k \xi_{nk}\right) \to \mathscr{L}(\xi), f_\xi(t) = e^{\psi(t)}, \sum\limits_k \sigma_{nk}^2 \to D\xi$ 的充分必要条件是 (8) 式中的 $\mu_n \xrightarrow{w} \mu$, 其中 μ 和 $\psi(t)$ 由 (5) 式相互唯一决定.

证 首先, 由引理 1 的证明知在条件 (C) 之下, 对固定的 t, 当 $n \geqslant n_t$ 时 $\sum\limits_k \log f_{nk}(t)$ 有意义且有界, 因而如果 $\prod\limits_k f_{nk}(t) \to g(t)$, 则由引理 1 及引理 1 之后 $\psi_n(t)$ 的定义知

$$\begin{aligned}\lim_{n\to\infty} e^{\psi_n(t)} &= \lim_{n\to\infty} \exp\left\{\sum_k (f_{nk}(t)-1)\right\} \\ &= \lim_{n\to\infty} \exp\left\{\sum_k \log f_{nk}(t)\right\} = g(t),\end{aligned}$$

再由引理 3 知必有一有限 L–S 测度 μ, 使得在 μ 的任一连续区间 I 上,$\mu_n(I) \to \mu(I)$ 且此时

$$g(t) = e^{\psi(t)},$$

其中

$$\psi(t)=\int\frac{e^{itx}-1-itx}{x^2}d\mu.$$

这样 i) 及 ii) 中条件的必要性获证.

再者, 若存在 μ 使得对 μ 的一切连续区间 I 均有 $\mu_n(I)\to\mu(I)$, 则 $\mu(R^{(1)})$ 有限, 由引理 3 知相应的 $e^{\psi_n(t)}\to e^{\psi(t)}$. 次由引理 1 知 $\prod\limits_k f_{nk}(t)=\exp\left\{\sum\limits_k\log f_{nk}(t)\right\}\to e^{\psi(t)}$, 再次由引理 4 知 $e^{\psi(t)}$ 为某一随机变量 ξ 的特征函数, 即 $\mathscr{L}\left(\sum\limits_k\xi_{nk}\right)\to\mathscr{L}(\xi), f_\xi(t)=e^{\psi(t)}$. 因而 ii) 中条件的充分性获证.

最后, 注意到 $\mu_n(R^{(1)})=\sum\limits_k\sigma_{nk}^2$ 及 $\mu(R^{(1)})=i^{-2}\frac{d^2}{dt^2}e^{\psi(t)}|_{t=0}=D\xi$, 再由 ii) 及 $\mu_n\xrightarrow{w}\mu$ 的定义知 iii) 成立. □

以上的讨论, 假定了所考虑的随机变量已中心化. 以下讨论非中心化的情况.

设 $\xi_{nk},k=1,\cdots,K_n$ 独立,$E\xi_{nk}=a_{nk},D\xi_{nk}=\sigma_{nk}^2$ 存在且有限, 并满足条件 (C), 令

$$\bar{P}_{nk}(B)=P_{\xi_{nk}-a_{nk}}(B),\bar{f}_{nk}=e^{-ia_{nk}t}f_{\xi_{nk}}(t); \tag{9}$$

$$\mu_n(B)=\sum_k\int_B x^2d\bar{P}_{nk}; \tag{10}$$

$$\psi_n(t)=ia_nt+\int\frac{e^{itx}-1-itx}{x^2}d\mu_n, a_n=\sum_k a_{nk}; \tag{11}$$

$$\psi(t)=iat+\int\frac{e^{itx}-1-itx}{x^2}d\mu,\mu(R)<\infty; \tag{12}$$

$$\bar{\psi}_n(t)=\psi_n(t)-ia_nt,\bar{\psi}(t)=\psi(t)-iat \tag{13}$$

则以 $\bar{f}_{nk}$ 代替 f_{nk}, 引理 1 成立.

引理 5　(12) 中的 ψ 与 (a,μ) 相互唯一决定.

证　(a,μ) 决定 ψ 显然, 反之, 由于

$$\psi'(t)=i\int\frac{e^{itx}-1}{x}d\mu+ia,\psi'(0)=ia.$$

$$\psi''(t)=-\int e^{itx}d\mu$$

因而 $\psi(t)$ 唯一决定 a 和 μ. □

引理 6　i)$e^{\psi_n(t)}\to$(某一)$g(t)$ 的充分必要条件是存在有限 L-S 测度 μ 及常数 a, 使得在 μ 的一切连续区间 I 上 $\mu_n(I)\to\mu(I)$ 且 $a_n\to a$.

ii) $e^{\psi_n(t)}$ 的极限函数为 $e^{\psi(t)},\psi(t)$ 具有 (12) 的形式.

证 条件的充分性显然. 今证必要性. 设 $e^{\psi_n(t)} \to g(t)$. 由于 $\{\mu_n\}$ 有界, 根据第 6 章 6.32 定理 4 知存在子序列 $\{\mu_{n'}\}$ 及有 限 L–S 测度 μ, 使对 μ 的一切连续区间 I 有 $\mu_{n'}(I) \to \mu(I)$. 再由 $\dfrac{e^{itx}-1-itx}{x^2}$ 对固定的 t 有界连续且当 $|x| \to \infty$ 时以零为极限, 故有 $\bar{\psi}_{n'}(t) \to \bar{\psi}(t)$. 因而

$$e^{ia_{n'}t} = e^{\psi_{n'}(t)-\bar{\psi}_{n'}(t)} \to g(t)e^{-\bar{\psi}(t)} (n \to \infty),$$

由第 6 章 §6.4 定理 3 知必存在 a, 使 $a_{n'} \to a$. 于是 $g(t) = e^{\psi(t)}, \psi(t)$ 必有 (12) 的形式. 最后由 ψ 与 (a, μ) 相互唯一决定知对自然数的任一子序列 $\{n''\}$, 都有 $\mu_{n''}(I) \to \mu(I)$ 且 $a_{n''} \to a$. 此即条件的必要性获证. □

引理 4 仍然成立, 因为 $e^{ita+\int \frac{e^{itx}-1-itx}{x^2}d\mu}$ 是特征函数为 $e^{\int \frac{e^{itx}-1-itx}{x^2}d\mu}$ 的随机变量的线性变换的特征函数.

应用这些引理可以证明

定理 2 设 $\xi_{nk}, k = 1, \cdots, k_n$ 是独立随机变量,$a_{nk} \triangleq E\xi_{nk}, \sigma_{nk}^2 \triangleq D\xi_{nk}$ 存在且有限, 并满足

$$\max_k \sigma_{nk}^2 \to 0, \sum_k \sigma_{nk}^2 \leqslant c < \infty, \tag{C}$$

则

i) 如果 $\prod\limits_k f_{nk}(t) \to g(t)$, 则 $g(t)$ 必是特征函数且具有 $e^{\psi(t)}$ 的形式, 其中 $\psi(t)$ 形如 (12). 从而满足条件 (C) 的独立随机变量和的一切可能的极限分布律的特征函数类与一切形如 $e^{\psi(t)}$ 的函数类重合.

ii) $\mathscr{L}\left(\sum\limits_k \xi_{nk}\right) \to \mathscr{L}(\xi)$ 且 $f_\xi(t) = e^{\psi(t)}$($\psi(t)$ 形如 (12), ξ 的特征函数必具此形式) 的充分必要条件是在 μ 的一切连续区间 I 上 $\mu_n(I) \to \mu(I)$ 且 $\sum\limits_k a_{nk} \to a$, 其中 μ_n 由 (10) 确定,ψ 与 (a, μ) 由 (12) 相互唯一确定.

iii) $\mathscr{L}\left(\sum\limits_k \xi_{nk}\right) \to \mathscr{L}(\xi), f_\xi(t) = e^{\psi(t)}$ 且 $\sum\limits_k \sigma_{nk}^2 \to D\xi$ 的充分必要条件是 (10) 中的 $\mu_n \xrightarrow{\text{w}} \mu$ 且 $\sum\limits_k a_{nk} \to a$. (其中 $\psi(t)$ 与 (a, μ) 由 (12) 相互唯一决定.)

证 只证 i), 至于 ii)iii) 的证明和定理 1 的证明类似, 不再赘述.

若 $\prod\limits_k f_{nk}(t) \to g(t)$, 即 $e^{i\sum\limits_k a_{nk}t}\prod\limits_k \bar{f}_{nk}(t) \to g(t)$. 由引理 1 知对每一固定的 t, 存在 n_t, 当 $n \geqslant n_t$ 时 $\sum\limits_k \log \bar{f}_{nk}(t)$ 有意义且有界,$\sum\limits_k \log \bar{f}_{nk}(t) - \sum\limits_k [\bar{f}_{nk}(t) - 1] \to 0$, 于是

$$\exp\left\{i\sum_k a_{nk}t+\sum_k \log \bar{f}_{nk}(t)\right\}\to g(t),$$
$$\exp\left\{i\sum_k a_{nk}t+\sum_k (\bar{f}_{nk}(t)-1)\right\}\to g(t).$$

即

$$\exp\{\psi_n(t)\}\to g(t).$$

于是由引理 3′ 知 $g(t)=e^{\psi(t)}, \psi(t)$ 具有 (12) 式定义的形式, 因而连续. 于是

$$\prod_k f_{nk}(t)\to e^{\psi(t)}.$$

$e^{\psi(t)}$ 连续, 因而是随机变量的特征函数. i) 获证. □

我们现在利用所得到的定理推出收敛于具体给定的分布律的条件.

1° 收敛于正态律.

正态律 $N(0,1)$ 的特征函数为 $e^{\frac{-t^2}{2}}$, 即 $\psi(t)=-\dfrac{t^2}{2}$. 故由 (5) 知与 $\psi(t)$ 相应的 μ 为

$$\mu(\{0\})=1, \mu(R^{(1)}-\{0\})=0.$$

由于唯一性引理 (引理 2), 只有上述的 μ 与 $N(0,1)$ 对应. 为了给出收敛于正态律的条件, 若 $E\xi_{nk}=0$, 只需给出 $\mu_n(I)\to\mu(I)$ 的条件 (其中 I 不以 0 点为端点), 即

$$\sum_k\int_{|x|\leqslant\varepsilon}x^2dP_{\xi_{nk}}\to 1, \varepsilon>0,$$
$$\sum_k\int_{\varepsilon_1<|x|<\varepsilon_2}x^2dP_{\xi_{nk}}\to 0, 0<\varepsilon_1<\varepsilon_2.$$

如果 $\sum\limits_k\sigma_{nk}^2\to 1\left(或\sum\limits_k\sigma_{nk}^2=1\right)$, 则必须且只需

$$\sum_k\int_{|x|\geqslant\varepsilon}x^2dP_{\xi_{nk}}\to 0, \varepsilon>0. \tag{14}$$

而后一条件成立时,

$$\max_k\sigma_{nk}^2=\max_k\int x^2dP_{\xi_{nk}}\leqslant\varepsilon^2+\max_k\int_{|x|\geqslant\varepsilon}x^2dP_{\xi_{nk}}$$

对任意 $\varepsilon>0$ 成立, 因而条件 (14) 蕴含着 $\max\sigma_{nk}^2\to 0$, 于是有

定理 3　设 $\xi_{nk}, k=1,\cdots,k_n$ 独立且 $E\xi_{nk}=0, \sigma_{nk}^2=D\xi_{nk}$ 有限, $k=1,\cdots,k_n, n=1,2,\cdots$ 且 $\sum\limits_k\sigma_{nk}^2=1\left(或\lim\limits_{n\to\infty}\sum\limits_k\sigma_{nk}^2=1\right)$. 则 $\mathscr{L}\left(\sum\limits_k\xi_{nk}\right)\to$

$N(0,1)$ 且 $\max_k \sigma_{nk}^2 \to 0$ 的充分必要条件是

$$g_n(\varepsilon) = \sum_k \int_{|x|>\varepsilon} x^2 dP_{nk} \to 0.$$

□

特别, 设 $\xi_1, \xi_2, \cdots$ 是一独立随机变量序列, $E\xi_n$ 和 $D\xi_n$ 存在有限, $n = 1, 2, \cdots$. 取

$$\xi_{nk} = \frac{\xi_k - E\xi_k}{S_n}, k = 1, \cdots, n, S_n^2 = \sum_{k=1}^n D\xi_k.$$

显然它们满足定理 3 的条件. 设 $F_k(x)$ 是 ξ_k 的分布函数, 易见 ξ_{nk} 的分布函数 $F_{nk}(x) = F_k(S_n x + E\xi_k)$, 于是

$$\begin{aligned} g_n(\varepsilon) &= \sum_{k=1}^n \int_{|x|\geqslant\varepsilon} x^2 dF_{nk}(x) \\ &= \sum_{k=1}^n \int_{|\frac{x-E\xi_k}{S_n}|\geqslant\varepsilon} |\frac{x - E\xi_k}{S_n}|^2 dF_k(x) \\ &= \frac{1}{S_n^2} \sum_{k=1}^n \int_{|x-E\xi_k|\geqslant\varepsilon S_n} |x - E\xi_k|^2 dF_k(x) \\ &\leqslant \frac{1}{\varepsilon^\delta S_n^{2+\delta}} \sum_{k=1}^n \int |x - E\xi_k|^{2+\delta} dF_k(x) \\ &= \frac{1}{\varepsilon^\delta S_n^{2+\delta}} \sum_{k=1}^n E|\xi_k - E\xi_k|^{2+\delta}, \end{aligned}$$

其中 δ 可以是任意的正数, 由此即得以下两个推论.

推论1(Lindeberg-Feller 定理) 设 $\{\xi_n\}$ 是独立随机变量序列, $E\xi_n$ 及 $D\xi_n = \sigma_n^2$ 存在, 则

$$\mathscr{L}\left(\frac{\sum_{k=1}^n(\xi_k - E\xi_k)}{S_n}\right) \to N(0,1) \text{及} \max_k \frac{\sigma_k}{S_n} \to 0$$

的充分与必要条件是对任一 $\varepsilon > 0$,

$$g_n(\varepsilon) = \frac{1}{S_n^2} \sum_{k=1}^n \int_{|x-E\xi_k|\geqslant\varepsilon S_n} |x - E\xi_k|^2 dP_{\xi_k} \to 0 (n \to \infty).$$

上述条件称为 Lindeberg 条件.

推论 2(Ляпунов 定理) 设 $\{\xi_n\}$ 是独立随机变量序列, 具有有限的数学期望 $E\xi_n$ 和方差 $D\xi_n$ 且对某一 $\delta > 0$

$$\frac{1}{S_n^{2+\delta}} \sum_{k=1}^n E|\xi - E\xi_k|^{2+\delta} \to 0, (n \to \infty).$$

则

$$\mathscr{L}\left(\frac{\sum_{k=1}^{n}(\xi_k - E\xi_k)}{S_n}\right) \to N(0,1).$$

2° 收敛于 Poisson 律. Poisson 律 $P(\lambda)$ 的特征函数为 $e^{\lambda(e^{it}-1)}$, 即 $\psi(t) = it\lambda + \lambda(e^{it} - 1 - it)$, 由此可见, 相应于定理 2 的 μ 为

$$\mu(\{1\}) = \lambda, \mu(R^{(1)} - \{1\}) = 0.$$

仿照上面 1° 的论述, 可得

定理 4　设 $\xi_{nk}, k = 1, \cdots, k_n$ 是独立随机变量, 具有有限的数学期望 $E\xi_{nk}$ 和方差 $D\xi_{nk} = \sigma_{nk}^2, k = 1, \cdots, k_n, n = 1, 2, \cdots$ 且 $\max\limits_k \sigma_{nk}^2 \to 0, \sum\limits_k \sigma_{nk}^2 \to \lambda$, 则 $\mathscr{L}\left(\sum\limits_k \xi_{nk}\right) \to P(\lambda)$ 的充分必要条件是 $\sum\limits_k E\xi_{nk} \to \lambda$ 且对任一 $\varepsilon > 0$,

$$\sum_k \int_{|x-1| \geqslant \varepsilon} x^2 dP_{\xi_{nk} - E\xi_{nk}} \to 0. \qquad \square$$

至于收敛向退化律, 可以看作收敛向正态律 $N(0,0)$ 的特殊情形而得到其判别准则.

§8.3　中心极限定理一般结果简介

在 §8.2 我们在有界方差情形下 (即 $\{\xi_{nk}\}$ 满足条件 (C)), 讨论了独立随机变量和 $\sum\limits_k \xi_{nk}$ 的极限分布问题. 对于一般情形, 即 $\{\xi_{nk}, k = 1, \cdots, k_n\}$ 独立, $n = 1, 2, \cdots$ 且满足一致可渐近忽略条件 (U) 的情形, 中心极限问题也已获得了完整的解答. 已经证明了一切可能的极限分布律是无穷可分律. 下面我们对无穷可分律的定义、条件 (U) 的等价形式、中心极限定理的一般结果及收敛向正态律、Poisson 律、退化律的判别条件作一简单介绍, 其证明请读者参看 [1] 或 [13].

定义　一个分布律称为**无穷可分律**, 如果它的特征函数 f 具有以下性质: 对于每一 n, 都存在一个特征函数 f_n 使 $f = f_n^n$. 这时我们也称特征函数 f**是无穷可分的**.

换句话说, 无穷可分律是这样的一种分布律, 对于任意 n, 总存在 n 个独立同分布的随机变量, 使得它刚好是这 n 个随机变量和的分布律.

许多常见的分布律都是无穷可分的. 例如退化律、正态律和 Poisson 律都是无穷可分的. 在第 2 节所述的一切形如 $e^{\psi(t)}$($\psi(t)$ 形如 (12)) 的特征函数都是无穷可

分的. 这是因为

$$e^{\psi(t)} = e^{iat+\int \frac{e^{itx}-1-itx}{x^2}d\mu(x)} = [e^{i\frac{a}{n}t+\int \frac{e^{itx}-1-itx}{x^2}d\mu_n(x)}]^n \triangleq [e^{\psi_n(t)}]^n,$$

其中 $\mu_n \triangleq \frac{1}{n}\mu, e^{\psi_n(t)}$ 仍然是特征函数.

定理 1 任何有限个无穷可分的特征函数的乘积仍是无穷可分的. 无穷可分律的极限分布仍是无穷可分律.

定理 2 每一无穷可分的特征函数均具有 $e^{\psi(t)}$ 的形式, 其中

$$\psi(t) = ita + \int \left(e^{itx} - 1 - \frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}d\mu(x), \tag{1}$$

其中 μ 为 $R^{(1)}$ 上有限 L–S 测度. 且 ψ 与 (a,μ) 相互唯一决定 (记作 $\psi = (a,\mu)$).

定理 2 相当于特征函数的唯一性定理, 而下面的定理 3 则相当于特征函数的连续性定理.

定理 3 如果 $a_n \to a, \mu_n \xrightarrow{w} \mu$, 则 $\psi_n \to \psi$. 反之, 若 $\psi_n \to g, g$ 在原点处连续, 则 $a_n \to a, \mu_n \xrightarrow{w} \mu$, 而 $g = \psi = (a,\mu)$.

定理 4(U 判据) 条件 (U) 与下列条件 (U_1) 或 (U_2) 分别等价:

$$\max_k \int \frac{x^2}{1+x^2}dP_{\xi_{nk}} \to 0, \tag{U_1}$$

$$\max_k |f_{\xi_{nk}}^{(t)} - 1| \to 0, \text{在}t\text{的任一有限区间上一致成立.} \tag{U_2}$$

定理 5 (中心极限定理) 设 $\xi_{nk}, k = 1, \cdots, k_n$ 是独立随机变量, 满足条件

$$\max_{1\leqslant k\leqslant k_n} P(|\xi_{nk}| \geqslant \varepsilon) \to 0(n \to \infty), \text{对一切}\varepsilon > 0. \tag{U}$$

则

1) 和 $\sum\limits_{k=1}^{k_n} \xi_{nk}$ 的一切可能的极限分布律是一切无穷可分律;

2) $\mathscr{L}\left(\sum\limits_k \xi_{nk}\right) \to \mathscr{L}(\xi), \xi$ 具有特征函数 $e^{\psi(t)}$, 其中

$$\psi(t) = ita + \int \left(e^{itx} - 1 - \frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}d\mu(x),$$

式中 $a \in R^{(1)}, \mu$ 是 $R^{(1)}$ 上有限 L–S 测度的充分必要条件是

$$a_n \to a, \quad \mu_n \xrightarrow{w} \mu,$$

其中

$$a_n = \sum_k \left\{ a_{nk} + \int \frac{x}{1+x^2} dP_{\xi_{nk}-a_{nk}} \right\},$$

$$\mu_n(A) = \sum_k \int_A \frac{x^2}{1+x^2} dP_{\xi_{nk}-a_{nk}}, A \in \mathscr{B}^{(1)}.$$

而

$$a_{nk} = \int_{|x|<\tau} x dP_{\xi_{nk}},$$

τ 是任意取定的一个正数.

定理 6 (推广的中心极限定理) 设 $\{\xi_{nk}, k = 1, 2, \cdots, k_n\}$ 是独立随机变量, $n = 1, 2, \cdots$ 满足条件 (U), 则

1) 和 $\sum\limits_k \xi_{nk} - a_n$ 的一切可能的极限分布律是一切无穷可分律,

2) 存在常数序列 a_n 使 $\mathscr{L}\left(\sum\limits_k \xi_{nk} - a_n\right)$ 收敛的充分必要条件是 $\mu_n \xrightarrow{w}$(某一)μ, 其中

$$\mu_n(A) = \sum_k \int_A \frac{x^2}{1+x^2} dP_{\xi_{nk}-a_{nk}},$$

而

$$a_{nk} = \int_{|x|<\tau} x dP_{\xi_{nk}}.$$

在收敛的情况下, 序列 $\{a_n\}$ 并不是唯一的, 一切可能的 a_n 都具有 $_n = \alpha_n - \alpha + o(1)$ 的形式. 其中 α 是任一有限常数

$$\alpha_n = \sum_k \left\{ a_{nk} + \int \frac{x}{1+x^2} dP_{\xi_{nk}-a_{nk}} \right\}$$

而相应的极限分布律的特征函数具有形式

$$e^{\psi(t)}, \psi = (\alpha, \mu).$$

定理 7(中心收敛判据) 设 $\xi_{nk}, k = 1, 2, \cdots, k_n$ 是独立随机变量, $n = 1, 2, \cdots$, 满足条件 (U), 则

$$\prod_k f_{nk} \to f = e^{\psi}, \psi = (\alpha, \mu)$$

的充分必要条件是

i) 对于一切使 $\mu(\{x\}) = 0$ 的 x,

$$\sum_k F_{\xi_{nk}}(x) \to \int_{-\infty}^{x} \frac{1+y^2}{y^2} d\mu, \text{当} x < 0.$$

$$\sum_k \{1 - F_{\xi_{nk}}(x)\} \to \int_x^\infty \frac{1+y^2}{y^2} d\mu, \text{当} x > 0.$$

ii) $\lim\limits_{\varepsilon \to 0} \overline{\lim\limits_{n\to\infty}} \sum\limits_k \left\{ \int_{|x|<\varepsilon} x^2 dP_{\xi_{nk}} - \left(\int_{|x|<\varepsilon} x dP_{\xi_{nk}} \right)^2 \right\} = \mu(\{0\})$,

其中 "$\lim\limits_{\varepsilon\to 0} \overline{\lim\limits_{n\to\infty}} \cdots = \cdots$" 表示 "$\lim\limits_{\varepsilon\to 0} \underline{\lim\limits_{n\to\infty}} \cdots = \lim\limits_{\varepsilon\to 0} \overline{\lim\limits_{n\to\infty}} \cdots = \cdots$".

iii) 对于一个固定的 $\tau > 0, [-\tau, \tau)$ 是 μ 的连续区间,

$$\sum_k \int_{|x|<\tau} x dP_{\xi_{nk}} \to \alpha + \int_{|x|<\tau} x d\mu - \int_{|x|\geqslant\tau} \frac{1}{x} d\mu.$$

最后我们介绍一下收敛向正态律、Poisson 律和退化律的判别准则.

我们用 $a_{nk}(\tau)$ 及 $\sigma_{nk}^2(\tau)$ 分别表示 ξ_{nk} 在 τ 处截割的随机变量 ξ_{nk}^τ 的数学期望和方差, 即

$$a_{nk}(\tau) = \int_{|x|<\tau} x dP_{\xi nk},$$

$$\sigma_{nk}^2(\tau) = \int_{|x|<\tau} x^2 dP_{\xi_{nk}} - \left(\int_{|x|<\tau} x dP_{\xi_{nk}} \right)^2.$$

1° 向正态律 $N(a, \sigma^2)$ 收敛.

$N(a,\sigma^2)$ 的特征函数对应的 $\psi(t) = (a, \mu), \mu(R^{(1)}) = \mu(\{0\}) = \sigma^2$.

定理 8 (正态收敛判据) 设 $\xi_{nk}, k = 1, \cdots, k_n$ 是独立随机变量, $n = 1, 2, \cdots$. 则

$$\mathscr{L}\left(\sum_k \xi_{nk} \right) \to N(a, \sigma^2), \max P(|\xi_{nk}| \geqslant \varepsilon) \to 0, (\text{对任一} \varepsilon > 0)$$

的充分必要条件是以下三条件同时成立.

$$\sum_k P(\sum |\xi_{nk}| \geqslant \varepsilon) \to 0, \text{对任一} \varepsilon > 0, \tag{N$_1$}$$

$$\sum_k \sigma_{nk}^2(\tau) \to \sigma^2, \text{对某一} \tau > 0, \tag{N$_2$}$$

$$\sum_k a_{nk}(\tau) \to a, \text{对某一} \tau > 0. \tag{N$_3$}$$

推论 1 设 $\xi_{nk}, k = 1, \cdots, k_n$ 是独立随机变量, $\mathscr{L}\left(\sum\limits_k \xi_{nk} \right)$ 收敛, 则极限律是正态的并且 $\{\xi_{nk}\}$ 满足条件 (U) 的充分必要条件是

$$\max_k |\xi_{nk}| \xrightarrow{P} 0.$$

2° 向 Poisson 律 $P(\lambda)$ 收敛.

Poisson 律 $P(\lambda)$ 的特征函数对应的 $\psi(t)=\left(\frac{\lambda}{2},\mu\right), \mu(R^{(1)})=\mu(\{1\})=\frac{\lambda}{2}$.

定理 9 (Poisson 律收敛判据) 设 $\xi_{nk}, k=1,2,\cdots,k_n$ 独立, $n=1,2,\cdots$ 满足条件 (U), 则 $\mathscr{L}\left(\sum_k \xi_{nk}\right)\to P(\tau)$ 的充分必要条件是以下三条件同时成立.

$$\sum_k \int_{\substack{|x|\geqslant\varepsilon\\|x-1|\geqslant\varepsilon}} dP_{\xi_{nk}}\to 0 \text{且} \sum_k \int_{|x-1|<\varepsilon} dP_{\xi_{nk}}\to\lambda, \text{对任一}\varepsilon\in(0,1), \tag{P$_1$}$$

$$\sum_k \sigma_{nk}^2(\tau)\to 0, \text{对于某一}\tau\in(0,1), \tag{P$_2$}$$

$$\sum_k a_{nk}(\tau)\to 0, \text{对于某一}\tau\in(0,1). \tag{P$_3$}$$

3° 向退化律 $\mathscr{L}(0)$ 收敛.

退化律 $\mathscr{L}(0)$ 可以看作特殊的正态律 $N(0,0)$, 我们有

定理 10 (退化收敛判据) 设 $\xi_{nk}, k=1,\cdots,k_n$ 独立, 则

$$\mathscr{L}\left(\sum_k \xi_{nk}\right)\to\mathscr{L}(0)$$

且 $\{\xi_{nk}\}$ 满足条件 (U) 的充分必要条件为对于每一 $\varepsilon>0$ 及一个 $\tau>0$,

$$\sum_k \int_{|x|\geqslant\varepsilon} dP\xi_{nk}\to 0, \tag{D$_1$}$$

$$\sum_k \sigma_{nk}^2(\tau)\to 0, \tag{D$_2$}$$

$$\sum_k a_{nk}(\tau)\to 0. \tag{D$_3$}$$

由此立得

推论 2 如果 $\{\xi_n\}$ 是独立随机变量序列, $b_n\uparrow\infty$, 则

$$\mathscr{L}\left(\frac{\sum\limits_{k=1}^{n}\xi_k}{b_n}\right)\to\mathscr{L}(0)$$

的充分必要条件是对每一 $\varepsilon>0$,

$$\sum_k \int_{|x|\geqslant\varepsilon b_n} dP_{\xi_k}\to 0,$$

$$\frac{1}{b_n^2}\sum_{k=1}^{n}\left\{\int_{|x|<b_n} x^2 dP_{\xi_k}-\left(\int_{|x|<b_n} x dP_{\xi_k}\right)^2\right\}\to 0,$$

$$\frac{1}{b_n}\sum_{k=1}^{n}\int_{|x|<b_n} x dP_{\xi_k}\to 0.$$

习题及补充

1. 设 $\xi_{nk}, k=1,\cdots,k_n$ 独立, $E\xi_{nk}=0, D\xi_{nk}=\sigma_{nk}^2$ 有限, $k=1,\cdots,k_n, n=1,2,\cdots$ 且满足条件 (C)

$$\max_k \sigma_{nk}^2\to 0, \sum_k \sigma_{nk}^2\leqslant c<\infty. \tag{C}$$

若 $\prod\limits_k f_{nk}(t)\xrightarrow{\text{a.e}} g(t)$, 则 $\sum\limits_k \xi_{nk}$ 依分布律收敛, 极限函数具有 $e^{\psi(t)}$ 的形式, 且 $g(t)=e^{\psi(t)}$, a.e.(L 测度), 其中 $\psi(t)$ 为

$$\psi(t)=\int\frac{e^{itx}-1-itx}{x^2}d\mu.$$

μ 是有限 L–S 测度.

2. 设 $\xi_1,\xi_2,\cdots$ 是独立随机变量序列, $S_n=\sum\limits_{k=1}^{n}\xi_k, s_n^2=\sum\limits_{k=1}^{n}D\xi_k$, 试判断在下述哪种情况下

$$\mathscr{L}\left(\frac{S_n-ES_n}{s_n}\right)\to N(0,1).$$

i) $P(\xi_k=-2^k)=P(\xi_k=2^k)=\dfrac{1}{2}, k=1,2,\cdots$;

ii) $P(\xi_k=2^k)=P(\xi_k=-2^k)=2^{-(2k+1)}$,
$P(\xi_k=0)=1-2^{-2k}, k=1,2,\cdots$;

iii) $P(\xi_k=k)=P(\xi_k=-k)=\dfrac{1}{2}k^{-\frac{1}{2}}$
$P(\xi_k=0)=1-k^{-\frac{1}{2}}, k=1,2,\cdots$.

3. 设 S_n 是 n 次独立试验中事件 A 出现的次数, p_k 为第 k 次试验 A 出现的概率. 试证

$$\lim_{n\to\infty}P\left(\frac{S_n-\sum\limits_{k=1}^{n}p_k}{\sqrt{\sum\limits_{k=1}^{n}p_k(1-p_k)}}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$$

的充分必要条件是 $\sum\limits_{k=1}^{\infty}p_k(1-p_k)=\infty$.

4. 设 $\xi_1, \xi_2 \cdots$ 独立, 且

$$P(\xi_n = -n^\alpha) = P(\xi_n = n^\alpha) = \frac{1}{2},$$

试证: 在 $\alpha > -\frac{1}{2}$ 时, Ляпунов定理的条件满足.

5. 求证

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

(提示: 应用Ляпунов定理到这样的独立随机变量之和上去: ξ_k 均按参数 $\lambda = 1$ 的 Poisson 律分布.)

6. 设 $\xi_1, \xi_2, \cdots$ 独立且一致有界, 则当

$$s_n^2 = \sum_{k=1}^{n} D\xi_k \to \infty$$

时,

$$\mathscr{L}\left(\frac{\sum_{k=1}^{n}(\xi_k - E\xi_k)}{s_n}\right) \to N(0, 1).$$

参考文献

[1] Loève M. Probability Theory. 4th edition. Springer-Verlag, 1977 (部分有中译本, 梁文骐译, 概率论, 上册, 科学出版社).

[2] Halmos P R. Measure Theory, Van Nostrand, 1950 (有中译本, 王建华译, 测度论, 科学出版社).

[3] 王梓坤. 概率论基础及其应用. 北京: 科学出版社, 1976.

[4] 王梓坤. 随机过程论. 北京: 科学出版社, 1965.

[5] Parthasarathy K. R. Introduction to Probability and Measure, Macmillan, 1977.

[6] Parthasarathy K R. Probability Measures on Metric Spaces, Academic Press, 1967.

[7] Neveu J. Bases Mathématiques du Calcul des Probabilités, Masson, 1964(有英译本, Mathematical Foundations of the Calculus of Probability, Holden-Day, 1965).

[8] Doob J L. Stochastic Processes, John Wiley, 1953.

[9] Cramer H. Mathematical Methods of Statistics, Princeton Univ. Press, 1946(有中译本, 魏宗舒等译, 统计学数学方法, 上海科学技术出版社).

[10] Wilks S S. Mathematical Statistics, John Wiley & Sons, 1962.

[11] Billingsley P. Convergence of Probability Measures, John Wiley & Sons, 1968.

[12] Гнеденко Б. В Курс Теории Вероятностей, Гостехиздат 1954 (有中译本, 丁寿田译, 概率论教程, 人民教育出版社).

[13] Гнеденко Б В и Колмогоров А Н. Предельные Распределения для Сумм Независимых Случайных Величин, Гостехиздат, 1949 (有中译本, 王寿仁译, 相互独立随机变数之和的极限分布, 科学出版社).

[14] Feller W. An Introduction to Probability Theory and its Applications, Vol. 1, 2d edition, 1975 (有中译本, 第一分册, 胡迪鹤, 林向清译; 第二分册, 刘文译, 概率论及其应用, 科学出版社).

[15] Lukacs E. Characteristic Functions, Charles Griffin, 1960.

[16] Дынкин Е Б. Основания Теория Марковских процессов, Физматгиз, 1959 (有中译本, 王梓坤译, 马尔科夫过程论基础, 科学出版社).

[17] Фихтенгольц Г М. Курс Дефференциального и Интегральиого Исчисления, Гостехиздат, 1951 (有中译本, 杨弢亮等译, 微积分学教程, 高等教育出版社).

[18] Привалов И И. Введение а Теорию Функцнй Комплексного Переменнного, Физматгиз, 1960 (有中译本, 闵嗣鹤等译, 复变函数引论, 人民教育出版社).

[19] 谢邦杰. 线性代数. 人民教育出版社, 1978.

[20] Dunford N and Schwartz J T. Linear Operators, Part I, Interscience Publishers, 1958.

[21] Натансон И П. Теория Функций Вещественной Переменно й, Гостехиздат, 1950(有中译本, 徐瑞云译, 实变函数论, 人民教育出版社).

[22] Хинчин А Я. Математические Методы Теории Массового Обслуживания, Изд. АН СССР, 1955(有中译本, 张里千, 殷涌泉译, 公用事业理论的数学方法, 科学出版社).

[23] 张禾瑞. 近世代数基础. 人民教育出版社, 1978.

[24] Колмогоров А Н и Фомин С В. Элементы Теории Функций и Функционального Анализа, М., Наука.1976.

[25] 关肇直. 泛函分析讲义. 科学出版社, 1958.

符号索引

---

① 1.2.2 表示第 1 章 1.2.2.

② 1.2.2.ex. 表示第 1 章 §1.2 习题.

内容索引

一至三画

十三画以上

《现代数学基础丛书》已出版书目

(按出版时间排序)

1　数理逻辑基础（上册）　1981.1　胡世华、陆钟万　著
2　紧黎曼曲面引论　1981.3　伍鸿熙、吕以辇、陈志华　著
3　组合论（上册）　1981.10　柯召、魏万迪　著
4　数理统计引论　1981.11　陈希孺　著
5　多元统计分析引论　1982.6　张尧庭、方开泰　著
6　概率论基础　1982.8　严士健、王隽骧、刘秀芳　著
7　数理逻辑基础（下册）　1982.8　胡世华、陆钟万　著
8　群构造（上册）　1982.11　张远达　著
9　有限群构造（下册）　1982.12　张远达　著
10　环与代数　1983.3　刘绍学　著
11　测度论基础　1983.9　朱成熹　著
12　分析概率论　1984.4　胡迪鹤　著
13　巴拿赫空间引论　1984.8　定光桂　著
14　微分方程定性理论　1985.5　张芷芬、丁同仁、黄文灶、董镇喜　著
15　傅里叶积分算子理论及其应用　1985.9　仇庆久等　编
16　辛几何引论　1986.3　J.柯歇尔、邹异明　著
17　概率论基础和随机过程　1986.6　王寿仁　著
18　算子代数　1986.6　李炳仁　著
19　线性偏微分算子引论（上册）　1986.8　齐民友　著
20　实用微分几何引论　1986.11　苏步青等　著
21　微分动力系统原理　1987.2　张筑生　著
22　线性代数群表示导论（上册）　1987.2　曹锡华等　著
23　模型论基础　1987.8　王世强　著
24　递归论　1987.11　莫绍揆　著
25　有限群导引(上册)　1987.12　徐明曜　著
26　组合论（下册）　1987.12　柯召、魏万迪　著
27　拟共形映射及其在黎曼曲面论中的应用　1988.1　李忠　著
28　代数体函数与常微分方程　1988.2　何育赞　著
29　同调代数　1988.2　周伯壎　著

30 近代调和分析方法及其应用 1988.6 韩永生 著
31 带有时滞的动力系统的稳定性 1989.10 秦元勋等 编著
32 代数拓扑与示性类 1989.11 马德森著 吴英青、段海鲍译
33 非线性发展方程 1989.12 李大潜、陈韵梅 著
34 反应扩散方程引论 1990.2 叶其孝等 著
35 仿微分算子引论 1990.2 陈恕行等 编
36 公理集合论导引 1991.1 张锦文 著
37 解析数论基础 1991.2 潘承洞等 著
38 拓扑群引论 1991.3 黎景辉、冯绪宁 著
39 二阶椭圆型方程与椭圆型方程组 1991.4 陈亚浙、吴兰成 著
40 黎曼曲面 1991.4 吕以辇、张学莲 著
41 线性偏微分算子引论（下册） 1992.1 齐民友 著
42 复变函数逼近论 1992.3 沈燮昌 著
43 Banach 代数 1992.11 李炳仁 著
44 随机点过程及其应用 1992.12 邓永录等 著
45 丢番图逼近引论 1993.4 朱尧辰等 著
46 线性微分方程的非线性扰动 1994.2 徐登洲、马如云 著
47 广义哈密顿系统理论及其应用 1994.12 李继彬、赵晓华、刘正荣 著
48 线性整数规划的数学基础 1995.2 马仲蕃 著
49 单复变函数论中的几个论题 1995.8 庄圻泰 著
50 复解析动力系统 1995.10 吕以辇 著
51 组合矩阵论 1996.3 柳柏濂 著
52 Banach 空间中的非线性逼近理论 1997.5 徐士英、李冲、杨文善 著
53 有限典型群子空间轨道生成的格 1997.6 万哲先、霍元极 著
54 实分析导论 1998.2 丁传松等 著
55 对称性分岔理论基础 1998.3 唐云 著
56 Gel'fond-Baker 方法在丢番图方程中的应用 1998.10 乐茂华 著
57 半群的 S-系理论 1999.2 刘仲奎 著
58 有限群导引(下册) 1999.5 徐明曜等 著
59 随机模型的密度演化方法 1999.6 史定华 著
60 非线性偏微分复方程 1999.6 闻国椿 著
61 复合算子理论 1999.8 徐宪民 著
62 离散鞅及其应用 1999.9 史及民 编著
63 调和分析及其在偏微分方程中的应用 1999.10 苗长兴 著

64　惯性流形与近似惯性流形　2000.1　戴正德、郭柏灵　著
65　数学规划导论　2000.6　徐增堃　著
66　拓扑空间中的反例　2000.6　汪林、杨富春　编著
67　拓扑空间论　2000.7　高国士　著
68　非经典数理逻辑与近似推理　2000.9　王国俊　著
69　序半群引论　2001.1　谢祥云　著
70　动力系统的定性与分支理论　2001.2　罗定军、张祥、董梅芳　编著
71　随机分析学基础（第二版）　2001.3　黄志远　著
72　非线性动力系统分析引论　2001.9　盛昭瀚、马军海　著
73　高斯过程的样本轨道性质　2001.11　林正炎、陆传荣、张立新　著
74　数组合地图论　2001.11　刘彦佩　著
75　光滑映射的奇点理论　2002.1　李养成　著
76　动力系统的周期解与分支理论　2002.4　韩茂安　著
77　神经动力学模型方法和应用　2002.4　阮炯、顾凡及、蔡志杰　编著
78　同调论——代数拓扑之一　2002.7　沈信耀　著
79　金兹堡-朗道方程　2002.8　郭柏灵等　著
80　排队论基础　2002.10　孙荣恒、李建平　著
81　算子代数上线性映射引论　2002.12　侯晋川、崔建莲　著
82　微分方法中的变分方法　2003.2　陆文端　著
83　周期小波及其应用　2003.3　彭思龙、李登峰、谌秋辉　著
84　集值分析　2003.8　李雷、吴从炘　著
85　数理逻辑引论与归结原理　2003.8　王国俊　著
86　强偏差定理与分析方法　2003.8　刘　文　著
87　椭圆与抛物型方程引论　2003.9　伍卓群、尹景学、王春朋　著
88　有限典型群子空间轨道生成的格（第二版）　2003.10　万哲先、霍元极　著
89　调和分析及其在偏微分方程中的应用（第二版）　2004.3　苗长兴　著
90　稳定性和单纯性理论　2004.6　史念东　著
91　发展方程数值计算方法　2004.6　黄明游　编著
92　传染病动力学的数学建模与研究　2004.8　马知恩、周义仓、王稳地、靳　祯　著
93　模李超代数　2004.9　张永正、刘文德　著
94　巴拿赫空间中算子广义逆理论及其应用　2005.1　王玉文　著
95　巴拿赫空间结构和算子理想　2005.3　钟怀杰　著
96　脉冲微分系统引论　2005.3　傅希林、闫宝强、刘衍胜　著
97　代数学中的Frobenius结构　2005.7　汪明义　著

98　生存数据统计分析　2005.12　王启华　著
99　数理逻辑引论与归结原理(第二版)　2006.3　王国俊　著
100　数据包络分析　2006.3　魏权龄　著
101　代数群引论　2006.9　黎景辉　陈志杰　赵春来　著
102　矩阵结合方案　2006.9　王仰贤　霍元极　麻常利　著
103　椭圆曲线公钥密码导引　2006.10　祝跃飞　张亚娟　著
104　凝固过程动力学与交界面稳定性引论 2006.12　徐鉴君 著
105　椭圆与超椭圆曲线公钥密码的理论与实现　2006.12　王学理　裴定一　著
106　非线性演化方程的稳定性与分歧　2007.4　马天　汪宁宏　著
107　正规族理论及其应用　2007.4　顾永兴　庞学诚　方明亮　著
108　组合网络理论　2007.5　徐俊明　著
109　矩阵的半张量积:理论与应用　2007.5　程代展　齐洪胜　著
110　鞅与 Banach 空间几何学　2007.5　刘培德　著
111　非线性常微分方程边值问题　2007.6　葛渭高　著
112　戴维-斯特瓦尔松方程　2007.5　戴正德　蒋慕蓉　李栋龙　著
113　广义哈密顿系统理论及其应用　2007.7　李继彬　赵晓华　刘正荣　著
114　Adams 谱序列和球面稳定同伦群　2007.7　林金坤　著
115　矩阵理论及其应用　2007.8　陈公宁　著
116　集值随机过程引论　2007.8　张文修　李寿梅　汪振鹏　高　勇　著
117　偏微分方程的调和分析方法　2008.1　苗长兴　张　波　著
118　拓扑动力系统概论　2008.1　叶向东　黄　文　邵　松　著
119　线性微分方程的非线性扰动(第二版)　2008.3　徐登洲　马如云　著
120　数组合地图论(第二版)　2008.3　刘彦佩　著
121　半群的 S-系理论(第二版)　2008.3　刘仲奎　乔虎生　著
122　巴拿赫空间引论(第二版)　2008.4　定光桂　著
123　拓扑空间论(第二版)　2008.4　高国士　著
124　非经典数理逻辑与近似推理(第二版)　2008.5　王国俊　著
125　非参数蒙特卡罗检验及其应用　2008.8　朱力行　许王莉　著
126　Camassa-Holm 方程　2008.8　郭柏灵　田立新　杨灵娥　殷朝阳　著
127　环与代数(第二版)　2009.1　刘绍学　郭晋云　朱　彬　韩　阳　著
128　泛函微分方程的相空间理论及应用　2009.4　王　克　范　猛　著
129　概率论基础(第二版)　2009.8　严士健　王隽骧　刘秀芳　著